B47

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C12N 15/12, C07K 14/47, 16/18, A61K 38/17, G01N 33/68 — A2

(11) International Publication Number: WO 00/31263

(43) International Publication Date: 2 June 2000 (02.06.00)

(21) International Application Number: PCT/US99/28013

(22) International Filing Date: 23 November 1999 (23.11.99)

(30) Priority Data:

60/109,592	23 November 1998 (23.11.98)	US
60/118,610	4 February 1999 (04.02.99)	US
60/127,990	6 April 1999 (06.04.99)	US

(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications

US	60/109,592 (CIP)
Filed on	23 November 1998 (23.11.98)
US	60/118,610 (CIP)
Filed on	4 February 1999 (04.02.99)
US	60/127,990 (CIP)
Filed on	6 April 1999 (06.04.99)

(71) Applicant (*for all designated States except US*): INCYTE PHARMACEUTICALS, INC. [US/US]; 3174 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and
(75) Inventors/Applicants (*for US only*): HILLMAN, Jennifer, L. [US/US]; 230 Monroe Drive, #12, Mountain View, CA 94040 (US). TANG, Y., Tom [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). BANDMAN, Olga [US/US]; 366 Anna Avenue, Mountain View, CA 94043 (US). LAL, Preeti [IN/US]; 2382 Lass Drive, Santa Clara, CA 95054 (US). YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). LU, Dyung, Aina, M. [US/US]; 55 Park Belmont Place, San Jose, CA 95136 (US). BAUGHN, Mariah, R. [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). YANG, Junming [CN/US]; 7136 Clarendon Street, San Jose, CA 95129 (US). AZIMZAI, Yalda [US/US]; 2045 Rock Springs Drive, Hayward, CA 94545 (US).

(74) Agents: BILLINGS, Lucy, J. et al.; Incyte Pharmaceuticals, Inc., 3174 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: GTPASE ASSOCIATED PROTEINS

(57) Abstract

The invention provides human GTPase associated proteins (GTPAP) and polynucleotides which identify and encode GTPAP. The invention also provides expression vectors, host cells, antibodies, agonist, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with expression of GTPAP.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

GTPASE ASSOCIATED PROTEINS

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of GTPase associated proteins and to the use of these sequences in the diagnosis, treatment, and prevention of cell proliferative, autoimmune/inflammatory, and immune system disorders.

BACKGROUND OF THE INVENTION

Guanine nucleotide binding proteins (GTP-binding proteins) participate in a wide range of regulatory functions in all eukaryotic cells, including metabolism, cellular growth, differentiation, signal transduction, cytoskeletal organization, and intracellular vesicle transport and secretion. In higher organisms they are involved in signaling that regulates such processes as the immune response (Aussel, C. et al (1988) J. Immunol. 140:215-220), apoptosis, differentiation, and cell proliferation including oncogenesis (Dhanasekaran, N. et al. (1998) Oncogene 17:1383-1394). Exchange of bound GDP for GTP followed by hydrolysis of GTP to GDP provides the energy that enables GTP-binding proteins to alter their conformation and interact with other cellular components. The superfamily of GTP-binding proteins consists of several families and may be grouped as translational factors, heterotrimeric GTP-binding proteins involved in transmembrane signaling processes (also called G-proteins), and low molecular weight GTP-binding proteins including the proto-oncogene Ras proteins and products of rab, rap, rho, rac, smg21, smg25, YPT, SEC4, and ARF genes, and tubulins (Kaziro, Y. et al. (1991) Ann. Rev. Biochem. 60:349-400). In all cases, the GTPase activity is regulated through interactions with other proteins.

GTP-binding proteins involved in protein biosynthesis include initiation factor 2 (IF-2), elongation factor 2 (EF-Tu), and elongation factor G (EF-G), observed in prokaryotes; and initiation factor 2 (eIF-2), elongation factor Iα (EF-Iα) and elongation factor 2 (EF-2) observed in eukaryotes (Kaziro, supra). IF-2 promotes the GTP-dependent binding of the tRNA to the small subunit of the ribosome, the step that initiates protein translation. Similarly, elongation factors promote the binding of tRNA and GTP and the displacement of GDP after hydrolysis as protein biosynthesis proceeds.

Heterotrimeric GTP-binding proteins are composed of 3 subunits (α, β and γ) which, in their inactive conformation, associate as a trimer at the inner face of the plasma membrane. G_α binds GDP or GTP and contains the GTPase activity. The $\beta\gamma$ complex enhances binding of G_α to a receptor. $G\gamma$ is necessary for the folding and activity of $G\beta$. (Neer, E.J. et al. (1994) Nature 371:297-300.) Multiple homologs of each subunit have been identified in mammalian tissues, and different combinations of subunits have specific functions and tissue specificities. (Spiegel, A.M. (1997) J.

Inher. Metab. Dis. 20:113-121.) G protein activity is triggered by seven-transmembrane cell surface receptors (G-protein coupled receptors) which respond to lipid analogs, amino acids and their derivatives, peptides, cytokines, and specialized stimuli such as light, taste, and odor. Activation of the receptor by its stimulus causes the replacement of the G protein-bound GDP with GTP. G_{α}-GTP dissociates from the receptor/βγ complex and each of these separated components can interact with and regulate downstream effectors. The signaling stops when G_{α} hydrolyzes its bound GTP to GDP and reassociates with the βγ complex (Neer, supra).

The alpha subunits of heterotrimeric G proteins can be divided into four distinct classes. The α-s class is sensitive to ADP-ribosylation by pertussis toxin which uncouples the receptor:G-protein interaction. This uncoupling blocks signal transduction to receptors that decrease cAMP levels which normally regulate ion channels and activate phospholipases. The inhibitory α-I class is also susceptible to modification by pertussis toxin which prevents α-I from lowering cAMP levels. Two novel classes of α subunits refractory to pertussis toxin modification are α-q, which activates phospholipase C, and α-12, which has sequence homology with the Drosophila gene concertina and may contribute to the regulation of embryonic development (Simon, M.I. (1991) Science 252:802-808).

The mammalian Gβ and Gγ subunits, each about 340 amino acids long, share more than 80% homology. The Gβ subunit (also called transducin) contains seven repeating units, each about 43 amino acids long. The activity of both subunits may be regulated by other proteins such as calmodulin and phosducin or the neural protein GAP 43 (D. Clapham and E. Neer, 1993, Nature 365:403-406). The β and γ subunits are tightly associated. The β subunit sequences are highly conserved between species, implying that they perform a fundamentally important role in the organization and function of G-protein linked systems (Van der Voorn L. (1992) Febs. Lett. 307 (2):131-134). They contain seven tandem repeats of the WD-repeat sequence motif, a motif found in many proteins with regulatory functions. WD-repeat proteins contain from four to eight copies of a loosely conserved repeat of approximately 40 amino acids which participates in protein-protein interactions. Mutations and variant expression of β transducin proteins are linked with various disorders. Mutations in LIS1, a subunit of the human platelet activating factor acetylhydrolase, cause Miller-Dieker lissencephaly. RACK1 binds activated protein kinase C, and RbAp48 binds retinoblastoma protein. CstF is required for polyadenylation of mammalian pre-mRNA in vitro and associates with subunits of cleavage-stimulating factor. Defects in the regulation of β-catenin contribute to the neoplastic transformation of human cells. The WD40 repeats of the human F-box protein βTrCP mediate binding to β-catenin, thus regulating the targeted degradation of β-catenin by

ubiquitin ligase (Neer, supra; Hart, M. et al (1999) Curr. Biol. 9:207-210). The γ subunit primary structures are more variable than those of the β subunits. They are often post-translationally modified by isoprenylation and carboxyl-methylation of a cysteine residue four amino acids from the C-terminus; this appears to be necessary for the interaction of the βγ subunit with the membrane and with other GTP-binding proteins. The βγ subunit has been shown to modulate the activity of isoforms of adenylyl cyclase, phospholipase C, and some ion channels. It is involved in receptor phosphorylation via specific kinases, and has been implicated in the p21ras-dependent activation of the MAP kinase cascade and the recognition of specific receptors by GTP-binding proteins. (Clapham and Neer, supra).

G-proteins interact with a variety of effectors including adenylyl cyclase (Clapham and Neer, supra). The signaling pathway mediated by cAMP is mitogenic in hormone-dependent endocrine tissues such as adrenal cortex, thyroid, ovary, pituitary, and testes. Cancers in these tissues have been related to a mutationally activated form of a $G\alpha_s$ known as the gsp (Gs protein) oncogene (Dhanasekaran, supra). Another effector is phosducin, a retinal phosphoprotein, which forms a specific complex with retinal Gβ and Gγ (Gβγ) and modulates the ability of Gβγ to interact with retinal Gα (Clapham and Neer, supra).

Irregularities in the GTP-binding protein signaling cascade may result in abnormal activation of leukocytes and lymphocytes, leading to the tissue damage and destruction seen in many inflammatory and autoimmune diseases such as rheumatoid arthritis, biliary cirrhosis, hemolytic anemia, lupus erythematosus, and thyroiditis. Abnormal cell proliferation, including cyclic AMP stimulation of brain, thyroid, adrenal, and gonadal tissue proliferation is regulated by G proteins. Mutations in Gα subunits have been found in growth-hormone-secreting pituitary somatotroph tumors, hyperfunctioning thyroid adenomas, and ovarian and adrenal neoplasms (Meij, J.T.A. (1996) Mol. Cell. Biochem. 157:31-38; Aussel, supra).

LMW GTP-binding proteins are GTPases which regulate cell growth, cell cycle control, protein secretion, and intracellular vesicle interaction. They consist of single polypeptides which, like the alpha subunit of the heterotrimeric GTP-binding proteins, are able to bind to and hydrolyze GTP, thus cycling between an inactive and an active state. LMW GTP-binding proteins respond to extracellular signals from receptors and activating proteins by transducing mitogenic signals involved in various cell functions. The binding and hydrolysis of GTP regulates the response of LMW GTP-binding proteins and acts as an energy source during this process (Bokoch, G. M. and Der, C. J. (1993) FASEB J. 7:750-759).

At least sixty members of the LMW GTP-binding protein superfamily have been identified

and are currently grouped into the ras, rho, arf, sar1, ran, and rab subfamilies. Activated ras genes were initially found in human cancers, and subsequent studies confirmed that ras function is critical in determining whether cells continue to grow or become differentiated. Ras1 and Ras2 proteins stimulate adenylate cyclase (Kaziro, supra), affecting a broad array of cellular processes. Stimulation of cell surface receptors activates Ras which, in turn, activates cytoplasmic kinases. These kinases translocate to the nucleus and activate key transcription factors that control gene expression and protein synthesis (Barbacid, M. (1987) Ann. Rev Biochem. 56:779-827, Treisman, R. (1994) Curr. Opin.Genet. Dev. 4:96-98). Other members of the LMW GTP-binding protein superfamily have roles in signal transduction that vary with the function of the activated genes and the locations of the GTP-binding proteins that initiate the activity. Rho GTP-binding proteins control signal transduction pathways that link growth factor receptors to actin polymerization, which is necessary for normal cellular growth and division. The rab, arf, and sar1 families of proteins control the translocation of vesicles to and from membranes for protein processing, localization, and secretion. Vesicle- and target- specific identifiers (v-SNAREs and t-SNAREs) bind to each other and dock the vesicle to the acceptor membrane. The budding process is regulated by the closely related ADP ribosylation factors (ARFs) and SAR proteins, while rab proteins allow assembly of SNARE complexes and may play a role in removal of defective complexes (J. Rothman and F. Wieland (1996) Science 272:227-234). Ran GTP-binding proteins are located in the nucleus of cells and have a key role in nuclear protein import, the control of DNA synthesis, and cell-cycle progression (Hall, A. (1990) Science 249:635-640; Barbacid, M. (1987) Ann. Rev Biochem. 56:779-827; Ktistakis, N. (1998) BioEssays 20:495-504; and Sasaki, T. and Takai, Y. (1998) Biochem. Biophys. Res. Commun. 245:641-645).

The cycling of LMW GTP-binding proteins between the GTP-bound active form and the GDP-bound inactive form is regulated by additional proteins. Guanosine nucleotide exchange factors (GEFs) increase the rate of nucleotide dissociation by several orders of magnitude, thus facilitating release of GDP and loading with GTP. The best characterized is the mammalian homologue of the Drosophila Son-of-Sevenless protein. Certain Ras-family proteins are also regulated by guanine nucleotide dissociation inhibitors (GDIs), which inhibit GDP dissociation. The intrinsic rate of GTP hydrolysis of the LMW GTP-binding proteins is typically very slow, but it can be stimulated by several orders of magnitude by GTPase-activating proteins (GAPs) (Geyer, M. and Wittinghofer, A. (1997) Curr. Opin. Struct. Biol. 7:786-792). Both GEF and GAP activity may be controlled in response to extracellular stimuli and modulated by accessory proteins such as RalBP1 and POB1. Mutant Ras-family proteins, which bind but can not hydrolyze GTP, are permanently activated, and cause cell proliferation or cancer, as do GEFs that inappropriately activate LMW GTP-binding proteins, such as the human oncogene NET1, a Rho-GEF (Drivas, G. T. et al. (1990) Mol. Cell. Biol.

10:1793-1798; Alberts, A. S. and Treisman, R. (1998) EMBO J. 14:4075-4085).

A novel group of GTP-binding proteins is the GTP1/OBG family, which are found in species ranging from bacteria to yeast to humans. These proteins contain characteristic GTP- binding motifs and are similar to one another but do not show sequence homology to other GTP-binding proteins. The exact functions of these proteins are as yet uncertain, but they have been shown to be important for regulation of cell differentiation and development (Okamoto, S. and Ochi, K. (1998). Mol. Microbiol 30:107-119; Sazaka, T. et al. (1992) Biochem. Biophys. Res. Commun. 189:363-370).

The discovery of new GTPase associated proteins and the polynucleotides encoding them satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of cell proliferative, autoimmune/inflammatory, and immune system disorders.

SUMMARY OF THE INVENTION

The invention features substantially purified polypeptides, GTPase associated proteins, referred to collectively as "GTPAP" and individually as "GTPAP-1," "GTPAP-2," "GTPAP-3," "GTPAP-4," "GTPAP-5," "GTPAP-6," "GTPAP-7," "GTPAP-8," "GTPAP-9," "GTPAP-10," "GTPAP-11," "GTPAP-12," "GTPAP-13," "GTPAP-14," "GTPAP-15," "GTPAP-16," "GTPAP-17," "GTPAP-18," "GTPAP-19," "GTPAP-20," "GTPAP-21," "GTPAP-22," "GTPAP-23," "GTPAP-24," "GTPAP-25," "GTPAP-26," "GTPAP-27," "GTPAP-28," and "GTPAP-29." In one aspect, the invention provides a substantially purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-29 and fragments thereof. The invention also includes a polypeptide comprising an amino acid sequence that differs by one or more conservative amino acid substitutions from an amino acid sequence selected from the group consisting of SEQ ID NO:1-29.

The invention further provides a substantially purified variant having at least 90% amino acid identity to at least one of the amino acid sequences selected from the group consisting of SEQ ID NO:1-29 and fragments thereof. The invention also provides an isolated and purified polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-29 and fragments thereof. The invention also includes an isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-29 and fragments thereof.

Additionally, the invention provides an isolated and purified polynucleotide which hybridizes under stringent conditions to the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-29 and fragments thereof. The

invention also provides an isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide encoding the polypeptide comprising the amino acid sequence selected from the group consisting of SEQ ID NO:1-29 and fragments thereof.

The invention also provides a method for detecting a polynucleotide in a sample containing nucleic acids, the method comprising the steps of: (a) hybridizing the complement of the polynucleotide sequence to at least one of the polynucleotides of the sample, thereby forming a hybridization complex; and (b) detecting the hybridization complex, wherein the presence of the hybridization complex correlates with the presence of a polynucleotide in the sample. In one aspect, the method further comprises amplifying the polynucleotide prior to hybridization.

The invention also provides an isolated and purified polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:30-58 and fragments thereof. The invention further provides an isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide sequence selected from the group consisting of SEQ ID NO:30-58 and fragments thereof. The invention also provides an isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:30-58 and fragments thereof.

The invention further provides an expression vector containing at least a fragment of the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-29. In another aspect, the expression vector is contained within a host cell.

The invention also provides a method for producing a polypeptide, the method comprising the steps of: (a) culturing the host cell containing an expression vector containing a polynucleotide of the invention under conditions suitable for the expression of the polypeptide; and (b) recovering the polypeptide from the host cell culture.

The invention also provides a pharmaceutical composition comprising a substantially purified polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:1-29 and fragments thereof, in conjunction with a suitable pharmaceutical carrier.

The invention further includes a purified antibody which binds to a polypeptide selected from the group consisting of SEQ ID NO:1-29 and fragments thereof. The invention also provides a purified agonist and a purified antagonist to the polypeptide.

The invention also provides a method for treating or preventing a disorder associated with decreased expression or activity of GTPAP, the method comprising administering to a subject in need of such treatment an effective amount of a pharmaceutical composition comprising a substantially

purified polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:1-29 and fragments thereof, in conjunction with a suitable pharmaceutical carrier.

The invention also provides a method for treating or preventing a disorder associated with increased expression or activity of GTPAP, the method comprising administering to a subject in need of such treatment an effective amount of an antagonist of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-29 and fragments thereof.

BRIEF DESCRIPTION OF THE TABLES

Table 1 shows polypeptide and nucleotide sequence identification numbers (SEQ ID NOs), clone identification numbers (clone IDs), cDNA libraries, and cDNA fragments used to assemble full-length sequences encoding GTPAP.

Table 2 shows features of each polypeptide sequence, including potential motifs, homologous sequences, and methods, algorithms, and searchable databases used for analysis of GTPAP.

Table 3 shows selected fragments of each nucleic acid sequence; the tissue-specific expression patterns of each nucleic acid sequence as determined by northern analysis; diseases, disorders, or conditions associated with these tissues; and the vector into which each cDNA was cloned.

Table 4 describes the tissues used to construct the cDNA libraries from which cDNA clones encoding GTPAP were isolated.

Table 5 shows the tools, programs, and algorithms used to analyze GTPAP, along with applicable descriptions, references, and threshold parameters.

DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same

meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

"GTPAP" refers to the amino acid sequences of substantially purified GTPAP obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and human, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which intensifies or mimics the biological activity of GTPAP. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of GTPAP either by directly interacting with GTPAP or by acting on components of the biological pathway in which GTPAP participates.

An "allelic variant" is an alternative form of the gene encoding GTPAP. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. A gene may have none, one, or many allelic variants of its naturally occurring form. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

"Altered" nucleic acid sequences encoding GTPAP include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as GTPAP or a polypeptide with at least one functional characteristic of GTPAP. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding GTPAP, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding GTPAP. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent GTPAP. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of GTPAP is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged

amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

The terms "amino acid" and "amino acid sequence" refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. Where "amino acid sequence" is recited to refer to an amino acid sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

"Amplification" relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

The term "antagonist" refers to a molecule which inhibits or attenuates the biological activity of GTPAP. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of GTPAP either by directly interacting with GTPAP or by acting on components of the biological pathway in which GTPAP participates.

The term "antibody" refers to intact immunoglobulin molecules as well as to fragments thereof, such as Fab, $F(ab')_2$, and Fv fragments, which are capable of binding an epitopic determinant. Antibodies that bind GTPAP polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (particular regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "antisense" refers to any composition containing a nucleic acid sequence which is

complementary to the "sense" strand of a specific nucleic acid sequence. Antisense molecules may be produced by any method including synthesis or transcription. Once introduced into a cell, the complementary nucleotides combine with natural sequences produced by the cell to form duplexes and to block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the designation "positive" or "plus" can refer to the sense strand.

The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" refers to the capability of the natural, recombinant, or synthetic GTPAP, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The terms "complementary" and "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence "5' A-G-T 3'" bonds to the complementary sequence "3' T-C-A 5'." Complementarity between two single-stranded molecules may be "partial," such that only some of the nucleic acids bind, or it may be "complete," such that total complementarity exists between the single stranded molecules. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands. This is of particular importance in amplification reactions, which depend upon binding between nucleic acid strands, and in the design and use of peptide nucleic acid (PNA) molecules.

A "composition comprising a given polynucleotide sequence" and a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding GTPAP or fragments of GTPAP may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been resequenced to resolve uncalled bases, extended using the XL-PCR kit (Perkin-Elmer, Norwalk CT) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from the overlapping sequences of one or more Incyte Clones and, in some cases, one or more public domain ESTs, using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (GCG, Madison WI). Some sequences have been both extended and assembled to produce the consensus sequence.

"Conservative amino acid substitutions" are those substitutions that, when made, least interfere with the properties of the original protein, i.e., the structure and especially the function of the

protein is conserved and not significantly changed by such substitutions. The table below shows amino acids which may be substituted for an original amino acid in a protein and which are regarded as conservative amino acid substitutions.

Original Residue	Conservative Substitution
Ala	Gly, Ser
Arg	His, Lys
Asn	Asp, Gln, His
Asp	Asn, Glu
Cys	Ala, Ser
Gln	Asn, Glu, His
Glu	Asp, Gln, His
Gly	Ala
His	Asn, Arg, Gln, Glu
Ile	Leu, Val
Leu	Ile, Val
Lys	Arg, Gln, Glu
Met	Leu, Ile
Phe	His, Met, Leu, Trp, Tyr
Ser	Cys, Thr
Thr	Ser, Val
Trp	Phe, Tyr
Tyr	His, Phe, Trp
Val	Ile, Leu, Thr

Conservative amino acid substitutions generally maintain (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a beta sheet or alpha helical conformation, (b) the charge or hydrophobicity of the molecule at the site of the substitution, and/or (c) the bulk of the side chain.

A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to the chemical modification of a polypeptide sequence, or a polynucleotide sequence. Chemical modifications of a polynucleotide sequence can include, for example, replacement of hydrogen by an alkyl, acyl, hydroxyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

A "fragment" is a unique portion of GTPAP or the polynucleotide encoding GTPAP which is identical in sequence to but shorter in length than the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from 5 to 1000 contiguous nucleotides or amino acid residues. A fragment

used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid residues in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50% of a polypeptide) as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present embodiments.

A fragment of SEQ ID NO:30-58 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:30-58, for example, as distinct from any other sequence in the same genome. A fragment of SEQ ID NO:30-58 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:30-58 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:30-58 and the region of SEQ ID NO:30-58 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-29 is encoded by a fragment of SEQ ID NO:30-58. A fragment of SEQ ID NO:1-29 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-29. For example, a fragment of SEQ ID NO:1-29 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-29. The precise length of a fragment of SEQ ID NO:1-29 and the region of SEQ ID NO:1-29 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

The term "similarity" refers to a degree of complementarity. There may be partial similarity or complete similarity. The word "identity" may substitute for the word "similarity." A partially complementary sequence that at least partially inhibits an identical sequence from hybridizing to a target nucleic acid is referred to as "substantially similar." The inhibition of hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (Southern or northern blot, solution hybridization, and the like) under conditions of reduced stringency. A substantially similar sequence or hybridization probe will compete for and inhibit the binding of a completely similar (identical) sequence to the target sequence under conditions of reduced stringency. This is not to say that conditions of reduced stringency are such that non-specific binding is permitted, as reduced stringency conditions require that the binding of two sequences to one another be a specific (i.e.. a selective) interaction. The absence of non-specific binding may be tested by the use of a second target sequence which lacks even a partial degree of complementarity (e.g., less than about 30% similarity or identity). In the absence of non-specific binding, the

substantially similar sequence or probe will not hybridize to the second non-complementary target sequence.

The phrases "percent identity" and "% identity," as applied to polynucleotide sequences, refer to the percentage of residue matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences.

Percent identity between polynucleotide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS 8:189-191. For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polynucleotide sequence pairs.

Alternatively, a suite of commonly used and freely available sequence comparison algorithms is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at http://www.ncbi.nlm.nih.gov/BLAST/. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at http://www.ncbi.nlm.nih.gov/gorf/bl2.html. The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Reward for match: 1

Penalty for mismatch: -2

Open Gap: 5 and Extension Gap: 2 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 11

Filter: on

Percent identity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous nucleotides. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures, or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes in a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

The phrases "percent identity" and "% identity," as applied to polypeptide sequences, refer to the percentage of residue matches between at least two polypeptide sequences aligned using a standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some alignment methods take into account conservative amino acid substitutions. Such conservative substitutions, explained in more detail above, generally preserve the hydrophobicity and acidity at the site of substitution, thus preserving the structure (and therefore function) of the polypeptide.

Percent identity between polypeptide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program (described and referenced above). For pairwise alignments of polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap penalty=3, window=5, and "diagonals saved"=5. The PAM250 matrix is selected as the default residue weight table. As with polynucleotide alignments, the percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polypeptide sequence pairs.

Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise comparison of two polypeptide sequences, one may use the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) with blastp set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Open Gap: 11 and Extension Gap: 1 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

Percent identity may be measured over the length of an entire defined polypeptide sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size, and which contain all of the elements required for stable mitotic chromosome segregation and maintenance.

The term "humanized antibody" refers to antibody molecules in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization" refers to the process by which a polynucleotide strand anneals with a complementary strand through base pairing under defined hybridization conditions. Specific hybridization is an indication that two nucleic acid sequences share a high degree of identity. Specific hybridization complexes form under permissive annealing conditions and remain hybridized after the "washing" step(s). The washing step(s) is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 μg/ml denatured salmon sperm DNA.

Generally, stringency of hybridization is expressed, in part, with reference to the temperature under which the wash step is carried out. Generally, such wash temperatures are selected to be about 5°C to 20°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. An equation for calculating T_m and conditions for nucleic acid hybridization are well known and can be found in Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY;

specifically see volume 2, chapter 9.

High stringency conditions for hybridization between polynucleotides of the present invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS, for 1 hour. Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%. Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, denatured salmon sperm DNA at about 100-200 μg/ml. Organic solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as for RNA:DNA hybridizations. Useful variations on these wash conditions will be readily apparent to those of ordinary skill in the art. Hybridization, particularly under high stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C_0t or R_0t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words "insertion" and "addition" refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively.

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

The term "microarray" refers to an arrangement of distinct polynucleotides on a substrate.

The terms "element" and "array element" in a microarray context, refer to hybridizable polynucleotides arranged on the surface of a substrate.

The term "modulate" refers to a change in the activity of GTPAP. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of GTPAP.

The phrases "nucleic acid" and "nucleic acid sequence" refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

"Operably linked" refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with the second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Generally, operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

"Probe" refers to nucleic acid sequences encoding GTPAP, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acid sequences. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes. "Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR).

Probes and primers as used in the present invention typically comprise at least 15 contiguous nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also be employed, such as probes and primers that comprise at least 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers may be considerably longer than these examples, and it is understood that any length supported by the specification, including the tables, figures, and Sequence Listing, may be used.

Methods for preparing and using probes and primers are described in the references, for example Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; Ausubel et al., 1987, Current Protocols in Molecular Biology, Greene Publ. Assoc. & Wiley-Intersciences, New York NY; Innis et al., 1990, PCR Protocols, A Guide to Methods and Applications, Academic Press, San Diego CA. PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5. 1991, Whitehead Institute for Biomedical Research, Cambridge MA).

Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to

5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

A "recombinant nucleic acid" is a sequence that is not naturally occurring or has a sequence that is made by an artificial combination of two or more otherwise separated segments of sequence. This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques such as those described in Sambrook, supra. The term recombinant includes nucleic acids that have been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently, a recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to transform a cell.

Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a vaccinia virus, that could be use to vaccinate a mammal wherein the recombinant nucleic acid is expressed, inducing a protective immunological response in the mammal.

The term "sample" is used in its broadest sense. A sample suspected of containing nucleic acids encoding GTPAP, or fragments thereof, or GTPAP itself, may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA,

RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" and "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, an antagonist, a small molecule, or any natural or synthetic binding composition. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide containing the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least about 60% free, preferably about 75% free, and most preferably about 90% free from other components with which they are naturally associated.

A "substitution" refers to the replacement of one or more amino acids or nucleotides by different amino acids or nucleotides, respectively.

"Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

"Transformation" describes a process by which exogenous DNA enters and changes a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term "transformed" cells includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95% or at least 98% or greater sequence identity over a certain defined length. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may

have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternate splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides generally will have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or at least 98% or greater sequence identity over a certain defined length of one of the polypeptides.

THE INVENTION

The invention is based on the discovery of new human GTPase associated proteins (GTPAP), the polynucleotides encoding GTPAP, and the use of these compositions for the diagnosis, treatment, or prevention of cell proliferative, autoimmune/inflammatory, and immune system disorders.

Table 1 lists the Incyte clones used to assemble full length nucleotide sequences encoding GTPAP. Columns 1 and 2 show the sequence identification numbers (SEQ ID NOs) of the polypeptide and nucleotide sequences, respectively. Column 3 shows the clone IDs of the Incyte clones in which nucleic acids encoding each GTPAP were identified, and column 4 shows the cDNA libraries from which these clones were isolated. Column 5 shows Incyte clones and their corresponding cDNA libraries. Clones for which cDNA libraries are not indicated were derived from pooled cDNA libraries. The Incyte clones in column 5 were used to assemble the consensus nucleotide sequence of each GTPAP and are useful as fragments in hybridization technologies.

The columns of Table 2 show various properties of each of the polypeptides of the invention: column 1 references the SEQ ID NO; column 2 shows the number of amino acid residues in each polypeptide; column 3 shows potential phosphorylation sites; column 4 shows potential glycosylation sites; column 5 shows the amino acid residues comprising signature sequences and motifs; column 6 shows homologous sequences as identified by BLAST analysis; and column 7 shows analytical

methods and in some cases, searchable databases to which the analytical methods were applied. The methods of column 7 were used to characterize each polypeptide through sequence homology and protein motifs.

The columns of Table 3 show the tissue-specificity and diseases, disorders, or conditions associated with nucleotide sequences encoding GTPAP. The first column of Table 3 lists the nucleotide SEQ ID NOs. Column 2 lists fragments of the nucleotide sequences of column 1. These fragments are useful, for example, in hybridization or amplification technologies to identify SEQ ID NO:30-58 and to distinguish between SEQ ID NO:30-58 and related polynucleotide sequences. The polypeptides encoded by these fragments are useful, for example, as immunogenic peptides. Column 3 lists tissue categories which express GTPAP as a fraction of total tissues expressing GTPAP. Column 4 lists diseases, disorders, or conditions associated with those tissues expressing GTPAP as a fraction of total tissues expressing GTPAP. Column 5 lists the vectors used to subclone each cDNA library. Of particular note is the specific expression of SEQ ID NO:43 in only one library, a human testis tissue library; the specific expression of SEQ ID NO:49 in only 4 libraries, one of which is associated with cell proliferation and 3 of which are associated with inflammation; and the specific expression of SEQ ID NO:40 in only 5 libraries, 3 of which are associated with cell proliferation and one of which is associated with inflammation.

The columns of Table 4 show descriptions of the tissues used to construct the cDNA libraries from which cDNA clones encoding GTPAP were isolated. Column 1 references the nucleotide SEQ ID NOs, column 2 shows the cDNA libraries from which these clones were isolated, and column 3 shows the tissue origins and other descriptive information relevant to the cDNA libraries in column 2.

The invention also encompasses GTPAP variants. A preferred GTPAP variant is one which has at least about 80%, or alternatively at least about 90%, or even at least about 95% amino acid sequence identity to the GTPAP amino acid sequence, and which contains at least one functional or structural characteristic of GTPAP.

The invention also encompasses polynucleotides which encode GTPAP. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:30-58, which encodes GTPAP.

The invention also encompasses a variant of a polynucleotide sequence encoding GTPAP. In particular, such a variant polynucleotide sequence will have at least about 70%, or alternatively at least about 90%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding GTPAP. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:30-58 which has at least about 70%, or alternatively at least about 90%, or even at least about

95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:30-58. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of GTPAP.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding GTPAP, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring GTPAP, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode GTPAP and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring GTPAP under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding GTPAP or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding GTPAP and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode GTPAP and GTPAP derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding GTPAP or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:30-58 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) Methods Enzymol. 152:399-407; Kimmel, A.R. (1987) Methods Enzymol. 152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment

of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Perkin-Elmer), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (Perkin-Elmer). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (Perkin-Elmer), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding GTPAP may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) PCR Methods Applic. 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 Primer Analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full-length cDNAs, it is preferable to use libraries that have been

size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Perkin-Elmer), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode GTPAP may be cloned in recombinant DNA molecules that direct expression of GTPAP, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express GTPAP.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter GTPAP-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

In another embodiment, sequences encoding GTPAP may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucleic Acids Symp. Ser. 7:215-223; and Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232.) Alternatively, GTPAP itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solid-phase techniques. (See, e.g., Roberge, J.Y. et al. (1995) Science 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (Perkin-Elmer). Additionally, the amino acid sequence of GTPAP, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g., Chiez, R.M. and F.Z. Regnier (1990) Methods Enzymol. 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY.)

In order to express a biologically active GTPAP, the nucleotide sequences encoding GTPAP or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding GTPAP. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding GTPAP. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding GTPAP and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) Results Probl. Cell Differ. 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding GTPAP and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding GTPAP. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or

tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g.. Ti or pBR322 plasmids); or animal cell systems. The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding GTPAP. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding GTPAP can be achieved using a multifunctional E. coli vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Life Technologies). Ligation of sequences encoding GTPAP into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of GTPAP are needed, e.g. for the production of antibodies, vectors which direct high level expression of GTPAP may be used. For example, vectors containing the strong, inducible T5 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of GTPAP. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, supra; Bitter, G.A. et al. (1987) Methods Enzymol. 153:516-544; and Scorer, C.A. et al. (1994) Bio/Technology 12:181-184.)

Plant systems may also be used for expression of GTPAP. Transcription of sequences encoding GTPAP may be driven viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J. 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See. e.g., Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843: and Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells. a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding GTPAP may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader

sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses GTPAP in host cells. (See, e.g., Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of GTPAP in cell lines is preferred. For example, sequences encoding GTPAP can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk*$^-$ and *apr*$^-$ cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), ß glucuronidase and its substrate ß-glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system.

(See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding GTPAP is inserted within a marker gene sequence, transformed cells containing sequences encoding GTPAP can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding GTPAP under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding GTPAP and that express GTPAP may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of GTPAP using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on GTPAP is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ.)

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding GTPAP include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding GTPAP, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for

ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding GTPAP may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode GTPAP may be designed to contain signal sequences which direct secretion of GTPAP through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding GTPAP may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric GTPAP protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of GTPAP activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the GTPAP encoding sequence and the heterologous protein sequence, so that GTPAP may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, supra, ch. 10).

A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled GTPAP may be achieved in vitro using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, for example, ^{35}S-methionine.

Fragments of GTPAP may be produced not only by recombinant means, but also by direct peptide synthesis using solid-phase techniques. (See, e.g., Creighton, supra, pp. 55-60.) Protein synthesis may be performed by manual techniques or by automation. Automated synthesis may be achieved, for example, using the ABI 431A peptide synthesizer (Perkin-Elmer). Various fragments of GTPAP may be synthesized separately and then combined to produce the full length molecule.

THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of GTPAP and GTPase associated proteins. In addition, the expression of GTPAP is closely associated with proliferating tissues associated with cancer and fetal development, inflamed tissues, and tissues invovled in the immune response. Therefore, GTPAP appears to play a role in cell proliferative, autoimmune/inflammatory, and immune system disorders. In the treatment of disorders associated with increased GTPAP expression or activity, it is desirable to decrease the expression or activity of GTPAP. In the treatment of disorders associated with decreased GTPAP expression or activity, it is desirable to increase the expression or activity of GTPAP.

Therefore, in one embodiment, GTPAP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of GTPAP. Examples of such disorders include, but are not limited to, a cell proliferative disorder, such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart. kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen. testis, thymus, thyroid, and uterus; an autoimmune/inflammatory disorder, such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies. ankylosing spondylitis. amyloidosis, anemia, asthma, atherosclerosis,

autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; and an immune system disorder, such as acquired immunodeficiency syndrome (AIDS), X-linked agammaglobinemia of Bruton, common variable immunodeficiency (CVI), DiGeorge's syndrome (thymic hypoplasia), thymic dysplasia, isolated IgA deficiency, severe combined immunodeficiency disease (SCID), immunodeficiency with thrombocytopenia and eczema (Wiskott-Aldrich syndrome), Chediak-Higashi syndrome, chronic granulomatous diseases, hereditary angioneurotic edema, and immunodeficiency associated with Cushing's disease, leukemias such as multiple myeloma, and lymphomas such as Hodgkin's disease.

In another embodiment, a vector capable of expressing GTPAP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of GTPAP including, but not limited to, those described above.

In a further embodiment, a pharmaceutical composition comprising a substantially purified GTPAP in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of GTPAP including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of GTPAP may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of GTPAP including, but not limited to, those listed above.

In a further embodiment, an antagonist of GTPAP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of GTPAP. Examples of such disorders include, but are not limited to, those cell proliferative, autoimmune/inflammatory, and immune system disorders described above. In one aspect, an antibody which specifically binds GTPAP may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for

bringing a pharmaceutical agent to cells or tissues which express GTPAP.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding GTPAP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of GTPAP including, but not limited to, those described above.

In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of GTPAP may be produced using methods which are generally known in the art. In particular, purified GTPAP may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind GTPAP. Antibodies to GTPAP may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are generally preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with GTPAP or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to GTPAP have an amino acid sequence consisting of at least about 5 amino acids, and generally will consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein and contain the entire amino acid sequence of a small, naturally occurring molecule. Short stretches of GTPAP amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to GTPAP may be prepared using any technique which provides for

the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. USA 80:2026-2030; and Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce GTPAP-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137.)

Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)

Antibody fragments which contain specific binding sites for GTPAP may also be generated. For example, such fragments include, but are not limited to, $F(ab')_2$ fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab')2 fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between GTPAP and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering GTPAP epitopes is generally used, but a competitive binding assay may also be employed (Pound, supra).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for GTPAP. Affinity is expressed as an

association constant, K_a, which is defined as the molar concentration of GTPAP-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple GTPAP epitopes, represents the average affinity, or avidity, of the antibodies for GTPAP. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular GTPAP epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10^9 to 10^{12} L/mole are preferred for use in immunoassays in which the GTPAP-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with K_a ranging from about 10^6 to 10^7 L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of GTPAP, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington, DC; Liddell, J.E. and Cryer, A. (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of GTPAP-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, supra, and Coligan et al. supra.)

In another embodiment of the invention, the polynucleotides encoding GTPAP, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, the complement of the polynucleotide encoding GTPAP may be used in situations in which it would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides encoding GTPAP. Thus, complementary molecules or fragments may be used to modulate GTPAP activity, or to achieve regulation of gene function. Such technology is now well known in the art, and sense or antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding GTPAP.

Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. Methods which are well known to those skilled in the art can be used to construct vectors to express nucleic acid sequences complementary to the polynucleotides

encoding GTPAP. (See, e.g., Sambrook, supra; Ausubel, 1995, supra.)

Genes encoding GTPAP can be turned off by transforming a cell or tissue with expression vectors which express high levels of a polynucleotide, or fragment thereof, encoding GTPAP. Such constructs may be used to introduce untranslatable sense or antisense sequences into a cell. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by endogenous nucleases. Transient expression may last for a month or more with a non-replicating vector, and may last even longer if appropriate replication elements are part of the vector system.

As mentioned above, modifications of gene expression can be obtained by designing complementary sequences or antisense molecules (DNA, RNA, or PNA) to the control, 5', or regulatory regions of the gene encoding GTPAP. Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may be employed. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding GTPAP.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques

for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding GTPAP. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nat. Biotechnol. 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as humans, dogs, cats, cows, horses, rabbits, and monkeys.

An additional embodiment of the invention relates to the administration of a pharmaceutical or sterile composition, in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic effects discussed above. Such pharmaceutical compositions may consist of GTPAP, antibodies to GTPAP, and mimetics, agonists, antagonists, or inhibitors of GTPAP. The compositions may be administered alone or in combination with at least one other agent, such as a stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone, or in combination with other agents, drugs, or hormones.

The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial,

intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA).

Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combining active compounds with solid excipient and processing the resultant mixture of granules (optionally, after grinding) to obtain tablets or dragee cores. Suitable auxiliaries can be added, if desired. Suitable excipients include carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, and sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums, including arabic and tragacanth; and proteins, such as gelatin and collagen. If desired, disintegrating or solubilizing agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, and alginic acid or a salt thereof, such as sodium alginate.

Dragee cores may be used in conjunction with suitable coatings, such as concentrated sugar solutions, which may also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, i.e., dosage.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with fillers or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's

solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils, such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate, triglycerides, or liposomes. Non-lipid polycationic amino polymers may also be used for delivery. Optionally, the suspension may also contain suitable stabilizers or agents to increase the solubility of the compounds and allow for the preparation of highly concentrated solutions.

For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention may be manufactured in a manner that is known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes.

The pharmaceutical composition may be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, and succinic acids. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preparation may be a lyophilized powder which may contain any or all of the following: 1 mM to 50 mM histidine, 0.1% to 2% sucrose, and 2% to 7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. For administration of GTPAP, such labeling would include amount, frequency, and method of administration.

Pharmaceutical compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example GTPAP or fragments thereof, antibodies of GTPAP, and agonists, antagonists or inhibitors of GTPAP, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be

determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED_{50} (the dose therapeutically effective in 50% of the population) or LD_{50} (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD_{50}/ED_{50} ratio. Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1 μg to 100,000 μg, up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

In another embodiment, antibodies which specifically bind GTPAP may be used for the diagnosis of disorders characterized by expression of GTPAP, or in assays to monitor patients being treated with GTPAP or agonists, antagonists, or inhibitors of GTPAP. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for GTPAP include methods which utilize the antibody and a label to detect GTPAP in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring GTPAP, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of GTPAP expression. Normal or standard values for GTPAP expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibody to GTPAP under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of GTPAP expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding GTPAP may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantify gene expression in biopsied tissues in which expression of GTPAP may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of GTPAP, and to monitor regulation of GTPAP levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding GTPAP or closely related molecules may be used to identify nucleic acid sequences which encode GTPAP. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification will determine whether the probe identifies only naturally occurring sequences encoding GTPAP, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the GTPAP encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:30-58 or from genomic sequences including promoters, enhancers, and introns of the GTPAP gene.

Means for producing specific hybridization probes for DNAs encoding GTPAP include the cloning of polynucleotide sequences encoding GTPAP or GTPAP derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ^{32}P or ^{35}S, or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding GTPAP may be used for the diagnosis of disorders associated with expression of GTPAP. Examples of such disorders include, but are not limited to, a cell proliferative disorder, such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an autoimmune/inflammatory disorder, such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; and an immune system disorder, such as acquired immunodeficiency syndrome (AIDS), X-linked agammaglobinemia of Bruton, common variable immunodeficiency (CVI), DiGeorge's syndrome (thymic hypoplasia), thymic dysplasia, isolated IgA deficiency, severe combined immunodeficiency disease (SCID), immunodeficiency with thrombocytopenia and eczema (Wiskott-Aldrich syndrome), Chediak-Higashi syndrome, chronic granulomatous diseases, hereditary angioneurotic edema, and immunodeficiency associated with Cushing's disease, leukemias such as multiple myeloma, and lymphomas such as Hodgkin's disease. The polynucleotide sequences encoding GTPAP may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered GTPAP expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding GTPAP may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding GTPAP may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding GTPAP in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of GTPAP, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding GTPAP, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding GTPAP may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced in vitro. Oligomers will preferably contain a fragment of a

polynucleotide encoding GTPAP, or a fragment of a polynucleotide complementary to the polynucleotide encoding GTPAP, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

Methods which may also be used to quantify the expression of GTPAP include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) J. Immunol. Methods 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as targets in a microarray. The microarray can be used to monitor the expression level of large numbers of genes simultaneously and to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, and to develop and monitor the activities of therapeutic agents.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.)

In another embodiment of the invention, nucleic acid sequences encoding GTPAP may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical chromosome mapping techniques and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) World Wide Web site. Correlation between the

location of the gene encoding GTPAP on a physical chromosomal map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder. The nucleotide sequences of the invention may be used to detect differences in gene sequences among normal, carrier, and affected individuals.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the disease or syndrome has been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, GTPAP, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between GTPAP and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with GTPAP, or fragments thereof, and washed. Bound GTPAP is then detected by methods well known in the art. Purified GTPAP can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding GTPAP specifically compete with a test compound for binding GTPAP. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with GTPAP.

In additional embodiments, the nucleotide sequences which encode GTPAP may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications, and publications mentioned above and below, in particular U.S. Ser. Nos. 60/109,592, 60/118,610, and 60/127,990 are hereby expressly incorporated by reference.

EXAMPLES

I. Construction of cDNA Libraries

RNA was purchased from Clontech or isolated from tissues described in Table 4. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A+) RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERSCRIPT plasmid system (Life Technologies), using the

recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, supra, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), or pINCY (Incyte Pharmaceuticals, Palo Alto CA). Recombinant plasmids were transformed into competent E. coli cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5α, DH10B, or ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

Plasmids were recovered from host cells by in vivo excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

III. Sequencing and Analysis

cDNA sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Perkin-Elmer) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer). Electrophoretic separation of cDNA sequencing reactions and detection of labeled

polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (Perkin-Elmer) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, supra, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example V.

The polynucleotide sequences derived from cDNA sequencing were assembled and analyzed using a combination of software programs which utilize algorithms well known to those skilled in the art. Table 5 summarizes the tools, programs, and algorithms used and provides applicable descriptions, references, and threshold parameters. The first column of Table 5 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score, the greater the homology between two sequences). Sequences were analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments were generated using the default parameters specified by the clustal algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

The polynucleotide sequences were validated by removing vector, linker, and polyA sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programing, and dinucleotide nearest neighbor analysis. The sequences were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM, and PFAM to acquire annotation using programs based on BLAST, FASTA, and BLIMPS. The sequences were assembled into full length polynucleotide sequences using programs based on Phred, Phrap, and Consed, and were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length amino acid sequences, and these full length sequences were subsequently analyzed by querying against databases such as the GenBank databases (described above), SwissProt, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, and Hidden Markov Model (HMM)-based protein family databases such as PFAM. HMM is a probabilistic approach which analyzes consensus primary structures of gene

families. (See, e.g., Eddy, S.R. (1996) Curr. Opin. Struct. Biol. 6:361-365.)

The programs described above for the assembly and analysis of full length polynucleotide and amino acid sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:30-58. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies were described in The Invention section above.

IV. Northern Analysis

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, supra, ch. 7; Ausubel, 1995, supra, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related molecules in nucleotide databases such as GenBank or LIFESEQ (Incyte Pharmaceuticals). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\%\ \text{sequence identity} \times \%\ \text{maximum BLAST score}}{100}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. For example, with a product score of 40, the match will be exact within a 1% to 2% error, and. with a product score of 70, the match will be exact. Similar molecules are usually identified by selecting those which show product scores between 15 and 40, although lower scores may identify related molecules.

The results of northern analyses are reported as a percentage distribution of libraries in which the transcript encoding GTPAP occurred. Analysis involved the categorization of cDNA libraries by organ/tissue and disease. The organ/tissue categories included cardiovascular, dermatologic, developmental, endocrine, gastrointestinal, hematopoietic/immune, musculoskeletal, nervous, reproductive, and urologic. The disease/condition categories included cancer, inflammation, trauma, cell proliferation, neurological, and pooled. For each category, the number of libraries expressing the sequence of interest was counted and divided by the total number of libraries across all categories. Percentage values of tissue-specific and disease- or condition-specific expression are reported in Table 3.

V. Extension of GTPAP Encoding Polynucleotides

The full length nucleic acid sequences of SEQ ID NO:30-58 were produced by extension of

an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer, to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg^{2+}, $(NH_4)_2SO_4$, and β-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 μl PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 μl of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 μl to 10 μl aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose mini-gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham

Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent E. coli cells. Transformed cells were selected on antibiotic-containing media, individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethysulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer).

In like manner, the nucleotide sequences of SEQ ID NO:30-58 are used to obtain 5' regulatory sequences using the procedure above, oligonucleotides designed for such extension, and an appropriate genomic library.

VI. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:30-58 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μCi of [γ-^{32}P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography or an alternative imaging means and

compared.

VII. Microarrays

A chemical coupling procedure and an ink jet device can be used to synthesize array elements on the surface of a substrate. (See, e.g., Baldeschweiler, supra.) An array analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced by hand or using available methods and machines and contain any appropriate number of elements. After hybridization, nonhybridized probes are removed and a scanner used to determine the levels and patterns of fluorescence. The degree of complementarity and the relative abundance of each probe which hybridizes to an element on the microarray may be assessed through analysis of the scanned images.

Full-length cDNAs, Expressed Sequence Tags (ESTs), or fragments thereof may comprise the elements of the microarray. Fragments suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). Full-length cDNAs, ESTs, or fragments thereof corresponding to one of the nucleotide sequences of the present invention, or selected at random from a cDNA library relevant to the present invention, are arranged on an appropriate substrate, e.g., a glass slide. The cDNA is fixed to the slide using, e.g., UV cross-linking followed by thermal and chemical treatments and subsequent drying. (See, e.g., Schena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645.) Fluorescent probes are prepared and used for hybridization to the elements on the substrate. The substrate is analyzed by procedures described above.

VIII. Complementary Polynucleotides

Sequences complementary to the GTPAP-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring GTPAP. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of GTPAP. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the GTPAP-encoding transcript.

IX. Expression of GTPAP

Expression and purification of GTPAP is achieved using bacterial or virus-based expression

systems. For expression of GTPAP in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*tac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express GTPAP upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of GTPAP in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding GTPAP by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, GTPAP is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from GTPAP at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, supra, ch. 10 and 16). Purified GTPAP obtained by these methods can be used directly in the following activity assay.

X. Demonstration of GTPAP Activity

The role of GTPAP can be assayed in vitro by monitoring the mobilization of Ca^{++} as part of the signal transduction pathway. (See, e.g., Grynkievicz, G. et al. (1985) J. Biol. Chem. 260:3440; McColl, S. et al. (1993) J. Immunol. 150:4550-4555; and Aussel, C. et al. (1988) J. Immunol. 140-215.) The assay requires preloading neutrophils or T cells with a fluorescent dye such as FURA-2.

Upon binding Ca^{++}, FURA-2 exhibits an absorption shift that can be observed by scanning the excitation spectrum between 300 and 400 nm, while monitoring the emission at 510 nm. When the cells are exposed to one or more activating stimuli artificially (i.e., anti-CD3 antibody ligation of the T cell receptor) or physiologically (i.e., by allogeneic stimulation), Ca^{++} flux takes place. Ca^{++} flux results from the release of Ca^{++} from intracellular organelles or from Ca^{++} entry into the cell through activated Ca^{++} channels. This flux can be observed and quantified by assaying the cells in a fluorometer or fluorescence activated cell sorter. Measurements of Ca^{++} flux are compared between cells in their normal state and those preloaded with GTPAP. Increased mobilization attributable to increased GTPAP availability results in increased emission.

Alternatively, GTPAP activity is measured by quantifying the amount of a non-hydrolyzable GTP analogue, GTPγS, bound over a 10 minute incubation period. Varying amounts of GTPAP are incubated at 30°C in 50mM Tris buffer, pH 7.5, containing 1mM dithiothreitol, 1mM EDTA and 1μM [^{35}S]GTPγS. Samples are passed through nitrocellulose filters and washed twice with a buffer consisting of 50mM Tris-HCl, pH 7.8, 1mM NaN_3, 10mM $MgCl_2$, 1mM EDTA, 0.5mM dithiothreitol, 0.01mM PMSF, and 200mM NaCl. The filter-bound counts are measured by liquid scintillation to quantify the amount of bound [^{35}S]GTPγS. GTPAP activity may also be measured as the amount of GTP hydrolysed over a 10 minute incubation period at 37°C. GTPAP is incubated in 50mM Tris-HCl buffer, pH 7.8, containing 1mM dithiothreitol, 2mM EDTA, 10μM [α-^{32}P]GTP, and 1μM H-rab protein. GTPase activity is initiated by adding $MgCl_2$ to a final concentration of 10 mM. Samples are removed at various time points, mixed with an equal volume of ice-cold 0.5mM EDTA, and frozen. Aliquots are spotted onto polyethyleneimine-cellulose thin layer chromatography plates, which are developed in 1M LiCl, dried, and autoradiographed. The signal detected is proportional to GTPAP activity.

Alternatively, GTPAP activity may be demonstrated as the ability to interact with its associated Gα or LMW GTPase in an <u>in vitro</u> binding assay. The candidate GTPases are expressed as fusion proteins with glutathione S-transferase (GST), and purified by affinity chromatography on glutathione-Sepharose. The GTPases are loaded with GDP by incubating 20 mM Tris buffer, pH 8.0, containing 100 mM NaCl, 2 mM EDTA, 5 mM MgCl2, 0.2 mM DTT, 100 μM AMP-PNP and 10 μM GDP at 30°C for 20 minutes. GTPAP is expressed as a FLAG fusion proteins in a baculovirus system. Extracts of these baculovirus cells containing GTPAP-FLAG fusion proteins are precleared with GST beads, then incubated with GST-GTPase fusion proteins. The complexes formed are precipitated by glutathione-Sepharose and separated by SDS-polyacrylamide gel electrophoresis. The separated proteins are blotted onto nitrocellulose membranes and probed with commercially available anti-

FLAG antibodies. GTPAP activity is proportional to the amount of GTPAP-FLAG fusion protein detected in the complex.

XI. Functional Assays

GTPAP function is assessed by expressing the sequences encoding GTPAP at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include pCMV SPORT (Life Technologies) and pCR3.1 (Invitrogen, Carlsbad CA), both of which contain the cytomegalovirus promoter. 5-10 μg of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2 μg of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of GTPAP on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding GTPAP and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding GTPAP and other genes of interest can be analyzed by northern analysis or microarray techniques.

XII. Producti n of GTPAP Specific Antibodies

GTPAP substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) Methods Enzymol. 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the GTPAP amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, supra, ch. 11.)

Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (Perkin-Elmer) using fmoc-chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, supra.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide and anti-GTPAP activity by, for example, binding the peptide or GTPAP to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XIII. Purification of Naturally Occurring GTPAP Using Specific Antibodies

Naturally occurring or recombinant GTPAP is substantially purified by immunoaffinity chromatography using antibodies specific for GTPAP. An immunoaffinity column is constructed by covalently coupling anti-GTPAP antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing GTPAP are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of GTPAP (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/GTPAP binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and GTPAP is collected.

XIV. Identification of Molecules Which Interact with GTPAP

GTPAP, or biologically active fragments thereof, are labeled with ^{125}I Bolton-Hunter reagent. (See, e.g., Bolton A.E. and W.M. Hunter (1973) Biochem. J. 133:529-539.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled GTPAP, washed, and any wells with labeled GTPAP complex are assayed. Data obtained using different concentrations of GTPAP are used to calculate values for the number, affinity, and association of

GTPAP with the candidate molecules.

Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

Table 1

Polypeptide SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
1	30	708398	SYNORAT04	568987X31 (MMLR3DT01), 708398H1, 708398X11, 708398X15, 708398X16, 708398X17, and 708398X21 (SYNORAT04), 2170523F6 (ENDCNOT03), 3374750H1 (CONNTUT05)
2	31	1259937	MENITUT03	913652R6 (STOMNOT02), 1259937F6 and 1259937H1 (MENITUT03), 1476721F1 (CORPNOT02), 1729248F6 (BRSTTUT08), 2191963H1 (THYRTUT03), 3129757F6 (LUNGTUT12), 3268746X15F1 (BRAINOT20), 3428891F6 (SKINNOT04)
3	32	1452285	PENITUT01	1452285F6 and 1452285H1 (PENITUT01), 2605011H1 (LUNGTUT07), 3505135H1 (ADRENOT11)
4	33	1812894	PROSTUT12	1812894H1, 1812894X12 and 1809113T6 (PROSTUT12), 1904479F6 (OVARNOT07), 2232535X15F1 and 2232535X18F1(PROSNOT16), 2267486X16C1 (UTRSNOT02), 2508562F6 (CONUTUT01)
5	34	3074884	BONEUNT01	225362F1 (PANCNOT01), 900707R1 (BRSTTUT03), 1339234F6 (COLNTUT03), 1759046R6 (PITUNOT03), 3074884H1 (BONEUNT01), SBDA02767F1
6	35	3452277	UTRSNON03	1684553F6 (PROSNOT15), 1951534H1 (PITUNOT01), 3452277H1 (UTRSNON03), 4092781T6 (BSCNSZT01), SBFA01413F1, SBFA03044F1, SBFA01805F1
7	36	4203832	BRAITUT29	723394F1 (SYNOOAT01), 862290R1, and 862290T1 (BRAITUT03), 1560918F1 (SPLNNOT04), 3509241H1 (CONCNOT01), 4203832H1 (BRAITUT29)
8	37	104368	BMARNOT02	104368H1 (BMARNOT02), SAEA03574F1, SAEA01063F1, SAEA00392F1, SAEA02287F1
9	38	1441680	THYRNOT03	1441680F6, 1441680H1, and 1441680T6 (THYRNOT03), 1904222F6 (OVARNOT07), 2477983F6 (SMCANOT01)
10	39	1494955	PROSNON01	965986R1 (BRSTNOT05), 1429037F1 and 1429037T1 (SINTBST01), 1453487F6 (PENITUT01), 1486114H1 (CORPNOT02), 1494955H1 (PROSNON01), 1995426R6 (BRSTTUT03), 2112074X18F1 and 2112348R6 (BRAITUT03)
11	40	1508161	LUNGNOT14	1508161F6 and 1508161H1 (LUNGNOT14), 3334303H1 (BRAIFET01), 4755656H1 (BRAHNOT01)
12	41	1811877	PROSTUT12	493795H1 (HNT2NOT01), 1573136H1 (LNODNOT03), 1811877F6 and 1811877H1 (PROSTUT12), 1825223F6 (LSUBNOT03), 2454143H1 (ENDANOT01), 2651022H1 (BLADTUT08), 3487062H1 (EPIGNOT01), 4536531H1 (OVARNOT12), 4795253H1 (LIVRTUT09), 4854087H1 (TESTNOT10), 4906149H2 (TLYMNOT08), 5196386H1 (LUNLTUT04)

Table 1 (cont.)

Polypeptide SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
13	42	1848674	LUNGFET03	1574127F6, 3857867X306F1, and 3857867X313F1 (LNODNOT03), 1848674H1 (LUNGFET03), 1877170F6 (LEUKNOT03), 2695307H1 (UTRSNOT12), 4148654H1 (SINITUT04), 4984182H1 (HELATXT05), 5288671H1 (LIVRTUS02)
14	43	2012970	TESTNOT03	2012970H1, 2012970R6, 2012970X11F (TESTNOT03)
15	44	2254315	OVARTUT01	022341F1 (ADENINB01), 198476R6 (KIDNNOT02), 2254315H1 (OVARTUT01), 2370170F6 (ADRENOT07), 2451278F6 (ENDANOT01)
16	45	2415545	HNT3AZT01	775722H1 (COLNNOT05), 870320R6 (LUNGAST01), 889023R1 (STOMTUT01), 895724R1 (BRSTNOT05), 1398541F1 (BRAITUT08), 1662585F6 (BRSTNOT09), 2415545H1 (HNT3AZT01), 2985066H1 (CARGDIT01), 3462702H1 (293TF2T01)
17	46	2707969	PONSAZT01	282552R1, 282552X23, and 282552X7 (CARDNOT01), 889783R1 (STOMTUT01), 1995451R6 (BRSTTUT03), 2707969H1 (PONSAZT01), SAAC00359R1.comp, SAAB00136R1, SAAC00330R1
18	47	2817769	BRSTNOT14	041660R1 (TBLYNOT01), 077378R1 (SYNORAB01), 740028R1 (PANCNOT04), 1593201F6 (BRAINOT14), 1924025R6 (BRSTTUT01), 2817769H1 (BRSTNOT14)
19	48	2917557	THYMFET03	473002F1 and 473002R1 (MMLR1DT01), 690999R6 (LUNGTUT02), 997483R1 (KIDNTUT01), 1430662F6 (SINTBST01), 1514017F1 (PANCTUT01), 1740475R6 (HIPONON01), 2109547H1 (BRAITUT03), 2917557H1 (THYMFET03), 4309528H1 (BRAUNOT01), 4990135H1 (LIVRTUT11)
20	49	3421335	UCMCNOT04	777588R6 and 777588T6 (COLNNOT05), 3421335H1 (UCMCNOT04)
21	50	605761	BRSTTUT01	605761F1, 605761H1, and 605761R6 (BRSTTUT01), 1271131X15 (TESTTUT02), 1516985F1 (PANCTUT01), 1524935H1 (UCMCL5T01), 2234846F6 (PANCTUT02)
22	51	483862	HNT2RAT01	483862H1 and 483862R1 (HNT2RAT01), 1750781X305F1, 1750781X307D2 (LIVRTUT01)
23	52	1256777	MENITUT03	264041R6 (HNT2AGT01), 826449R1 (PROSNOT06), 1256777H1 (MENITUT03), 2276061R6 (PROSNON01), 4614049H1 (BRAHNOT01)
24	53	2198779	SPLNFET02	1557708F6 (BLADTUT04), 1922490R6 (BRSTTUT01), 2198779H1 (SPLNFET02), 2541193F7 (BONRTUT01), 3039254F6 (BRSTNOT16), 3057079H1 (LNODNOT08), 3105017H1 (COLNUCT03), 4239592H1 (SYNWDIT01), 5064513H1 (ARTFTDT01)

Table 1 (cont.)

Polypeptide SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
25	54	2226116	SEMVNOT01	1662607F6 (BRSTNOT09), 1662607T6 (BRSTNOT09), 2226116F6 (SEMVNOT01), 2226116H1 (SEMVNOT01), 2930011F6 (TLYMNOT04), 3015747T6 (MUSCNOT07), 4087670H1 (LIVRNOT06)
26	55	2504472	CONUTUT01	420365F1 (BRSTNOT01), 762246R1 (BRAITUT02), 907754R2 (COLNNOT09), 1007508H1 (HEALDIT02), 1302342F6 (PLACNOT02), 1913887H1 (PROSTUT04), 2023822F6 (CONNNOT01), 2023822X11R1 (CONNNOT01), 2504472H1 (CONUTUT01), 2951618F6 (KIDNFET01)
27	56	3029920	HEARFET02	354846T6 (RATRNOT01), 418533R6 (BRSTNOT01), 935073R1 (CERVNOT01), 1340722F1 (COLNTUT03), 1416203T6 (BRAINOT12), 1524567F1 (UCMCL5T01), 1773043H1 (MENTUNON3), 2590310H2 (LUNGNOT22), 3029920H1 (HEARFET02), 4873053H1 (COLDNOT01), 5687696H1 (BRAIUNT01)
28	57	3332415	BRAIFET01	118166R1 (MUSCNOT01), 1257348H1 (MENITUT03), 1288237T6 (BRAINOT11), 1335936F6 (COLNNOT13), 1452268H1 (PENITUT01), 1996016R6 (BRSTTUT03), 2116665R6 (BRSTTUT02), 2206894F6 (SINTFET03), 2540063H1 (BONRTUT01), 2808268H1 (BLADTUT08), 3086221H1 (HEAONOT03), 3127508H1 (LUNGTUT12), 3295812H1 (TLYJINT01), 3332415H1 (BRAIFET01), 3604705H1 (LUNGNOT30), 4821203H1 (PROSTUT17), 4970353H1 (KIDEUNC10), 5055775H1 (COLATMT01)
29	58	4031536	BRAINOT23	029167X3 (SPLNFET01), 350137R1 (LVENNOT01), 408825X1 (EOSIHET02), 689446X23 (LUNGTUT02), 1963062R6 (BRSTNOT04), 2288043R6 (BRAINON01), 4031536H1 (BRAINOT23)

Table 2

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods and Databases
1	1002	T30 S224 T405 S499 T533 S558 S701 T737 T845 S864 S6 T152 T268 T412 T442 T464 T514 T528 T693 S814 S815 S823 T880 Y117 Y842	N446	G524-T531: ATP/GTP-binding site motif	GTP-binding protein [Mus musculus] g53169	BLAST MOTIFS
2	338	S21 S77 T86 S200 T246 T299 S77 S306 Y131	N244		cAMP-regulated Guanine nucleotide exchange factor [Rattus norvegicus] g4079657	BLAST
3	211	S159 S199	N33 N74	G16-T23: ATP/GTP-binding site motif	GTP-binding protein [Rattus norvegicus] g206543	BLAST MOTIFS PFAM BLOCKS PRINTS
4	516	T14 S42 T237 S270 S347 S360 T371 T395 T433 S500 T3 S13 S96 T316 S430			Fos-related antigen [Rattus norvegicus] g1016712 Rabaptin-4 [H. sapiens] g3832516	BLAST MOTIFS
5	445	T44 T114 T219 T297 S314 S341 S356 T412 T24 S72 T91 T328 T388 T394		G230-T237: ATP/GTP-binding site motif	GTP-binding protein [H. sapiens] g2765411	BLAST MOTIFS

Table 2 (cont.)

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods and Databases
6	445	S174 S202 S289 S29 S305 S323 T434 T11 T147 T197 T198 S270 S273 S371 S397 Y125	N73		Regulator of G-protein signaling-9 [H. sapiens] g3284012	BLAST
7	281	S182 S210 S254 S13 T56 S110 S182 S32 T46 S66 S177	N130 N181	G31-T38:ATP/GTP-binding site motif	Putative ras-like protein [H. sapiens] g4092830	MOTIFS PRINTS BLAST PFAM
8	301	S92 T2 T3 Y15 S18 S19 S20 S25 S97 T120 S165 S296 T94 S116 T120 S284		E47-G66, S116-E178, Y188-G272: Phosducin signature	Phosducin-like protein [Rattus rattus] g1323727	MOTIFS BLAST PRINTS
9	485	T6 Y57 S82 T91 S112 S187 T231 T257 S309 T6 T81 S132 S157 S210 S241 T462	N460	L49-S82: Beta G protein	Similar to WD domain Beta transducin-like protein [C. elegans] g5596646	MOTIFS BLAST PRINTS
10	447	S420 S94 T107 S118 T167 T179 T308 S390 S39 S58 T78 T113 S129 T160 T167 Y174 T199 S216 S291 T302 T323 T359 T384 S423 T438	N76 N92 N231 N289 N378 N421	M294-T308: Beta transducin	WS beta-transducin repeat protein [Homo sapiens] g4704417	MOTIFS BLAST
11	199	S90 T55 T140 S190		K6-E130: Ras Guanine exchange factor	Putative guanine nucleotide releasing factor [Drosophila affinis] g2981229	MOTIFS BLAST PFAM

Table 2 (cont.)

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods and Databases
12	694	S57 S67 S99 T150 T346 S416 S467 S500 T522 T684 S99 T156 S209 S285 T331 T360 T388 T430 T477 T650 T688		L10-I24, M96-L110: Beta transducin	Transducin-like protein [H. sapiens] g414536	MOTIFS BLAST
13	654	T10 S15 T49 S97 S102 S104 S112 S113 S377 S432 S638 T46 S54 S84 S97 T177 S217 T307 S401 S450 S504 T515 S546 T547 S561 Y618	N353 N362 N502	L197-F211: Beta transducin	Similar to the beta transducin family [C. elegans] g2315521	MOTIFS BLAST
14	180	S14		G23-S30: ATP-GTP binding site	Rab7C (small GTP binding protein) [Lotus japonicus] g1370186	MOTIFS BLAST
15	374	T100 T249 S260 T308 T328 S338 S351 S30 T73 T157 S237 T308	N114 N189 N222	G26-T33: ATP-GTP binding site	ATP(GTP)-binding protein [H. sapiens] g3646130	MOTIFS BLAST
16	649	S67 T344 S366 S63 S68 S75 S122 S177 S265 T282 T332 S373 S380 S563 T569 S634 S20 T94 S128 S314 T382 T385 T458 T559		F307-S544: Probable rabGAP domain	Similar to probable rabGAP [C. elegans] g3925265	MOTIFS BLAST PFAM
17	698	T244 S262 S17 T41 T42 T196 S206 S317 S479 S522 S556 T586 T680 T31 S95 T99 T140 T173 S257 T322 S374 T450 S568 T619	N171 N194 N685		Small GTP-binding protein associated protein [Mus musculus] g725274	MOTIFS BLAST

Table 2 (cont.)

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods and Databases
18	396	T325 S115 T133 S232 S275 T336 S22 T221 S232 T320	N60 N230 N286	G29-S36: ATP-GTP binding site	Putative GTP-binding protein [C. elegans] g3880615	MOTIFS BLAST
19	634	T197 S3 S5 S9 T14 S132 T197 T285 T553 T40 T56 S160 T189 S261 S582 Y20 Y396 Y419		G52-T59: ATP-GTP binding site	Putative GTP-binding protein [H. sapiens] g3169010	MOTIFS BLAST
20	196	T60 S73 S90 S99 S73 S193		G19-T26: ATP-GTP binding site	Kidney injury associated protein HW052 Acc No W86322 ADP-ribosylation factor-like protein 3 [Rattus norvegicus] g560006	MOTIFS BLAST
21	446	T10 T24 T93 S122 T243 S263 S270 T305 S317 S325 T357 S372 T379 S100 S170 S223 T227 S285 T348	N79	L323-L337: Beta transducin	Putative WD40 repeat protein [A. thaliana] g4191784	MOTIFS BLAST
22	265	T184 T76 T137 S139 T161 T174 T183 S213	N159	L141, L148, L155 L: zipper gene regulatory motif	TipD; similar to beta transducin family [D. discoideum] g2407788	MOTIFS BLAST

Table 2 (cont.)

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods and Databases
23	185	T55 S111 S127 S148 S171 S14 S94 Y103		G10-T17: ATP/GTP binding site (P-loop) A4-S72: Ras domain		MOTIFS PFAM PRINTS
24	554	S388 T488 S30 S75 T111 S149 S220 S237 T255 S305 S325 T339 T359 S363 S509 S172 T195 S211 T378 T438 T470 Y203	N5	N297-D336, P345-D383, G481-Q519: Beta-transducin WD40 repeats	WD-repeat protein [Arabidopsis thaliana] g3924603	BLAST MOTIFS PFAM PRINTS
25	434	S164 S341 T347 S36 S68 S92 T286 S364	N22 N383	G259-S266:ATP/GTP binding site (P-loop): G113-R433: GTP1/OBG domain	Predicted GTP binding protein [C. elegans] g3878629	BLAST MOTIFS PFAM BLOCKS PRINTS
26	826	S122 T243 T247 T427 S454 S519 T528 S623 S701 S715 S809 T58 S143 S266 T411 S505 S577 S603 T661 S735 T753 S791 T815	N23 N264 N576 N600 N789	R48-E91, L97-S143, F197 K237, V273-W319, W378-A416, W604 K642, A659-G697: Beta-transducin WD40 repeats	Predicted WD repeat protein [S. cerevesiae] P42935	BLAST MOTIFS PFAM PRINTS
27	618	T414 S59 T105 S126 T139 T143 S196 T203 S311 S325 T370 T390 S477 T483 S541 T583 T94 S148 T247 Y160 Y383 Y456	N118 N154 N346	G11-T18, G425-S432: ATP/GTP binding site (P-loop) R6-K187: Ras domain	GTP-binding protein APD08 [H.sapiens] Accession W75771	BLAST MOTIFS PFAM PRINTS

Table 2 (cont.)

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods and Databases
28	596	S17 S21 S50 S152 S153 T533 S539 T594 S36 S38 S80 T163 T169 S183 S211 T240 S306 T329 T417 S457 S508 T545 S45 T64 S88 T124 S139 S299 S451 S459 S528 S568 Y180 Y364		A178-L355: Rho-family guanine nucleotide exchange factor (RhoGEF) domain	Guanine nucleotide regulatory protein (NET1 homologue) [Mus musculus] g3834631	BLAST MOTIFS PFAM BLOCKS
29	589	T108 S20 T90 S127 S176 S278 S467 T521 S522 T189 S254 T284 T292 T321 T324 T345 T364 T423 S444 T484 T527	N572	L252-S289, G293-N329, G333-D369, G373-D409, E413-D449, G453-D489, G493-D532: Beta-transducin WD40 repeats R160-K206: F-box domain	SEL-10 [C.elegans] g2677836	BLAST MOTIFS PFAM PRINTS

Table 3

Nucleotide Seq ID NO:	Selected Fragments	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
30	628-711	Reproductive (0.256) Nervous (0.154) Gastrointestinal (0.154)	Cell Proliferation (0.692) Inflammation (0.372)	PSPORT1
31	1094-1129	Reproductive (0.268) Cardiovascular (0.146) Nervous (0.146)	Cell Proliferation (0.731) Inflammation (0.219) Neurological (0.049)	pINCY
32	652-703	Cardiovascular (0.375) Reproductive (0.375) Dermatologic (0.125) Endocrine (0.125)	Cell Proliferation (0.875) Trauma (0.125)	pINCY
33	1224-1292	Reproductive (0.412) Gastrointestinal (0.147) Hematopoietic/Immune (0.147)	Cell Proliferation (0.647) Inflammation (0.264)	pINCY
34	16-65	Nervous (0.211) Reproductive (0.197) Gastrointestinal (0.169)	Cell Proliferation (0.507) Inflammation (0.352)	pINCY
35	947-1043	Reproductive (0.444) Nervous (0.333) Gastrointestinal (0.111) Urologic (0.111)	Cell Proliferation (0.667) Inflammation (0.111) Neurological (0.111)	pINCY
36	840-1001	Nervous (0.340) Reproductive (0.208) Gastrointestinal (0.151)	Cell Proliferation (0.641) Inflammation (0.302) Neurological (0.038)	pINCY

Table 3 (cont.)

Nucleotide Seq ID NO:	Selected Fragments	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
37	507-551	Hematopoietic/Immune (0.269) Nervous (0.269) Reproductive (0.154)	Inflammation (0.423) Cell Proliferation (0.269)	PBLUESCRIPT
38	218-262	Cardiovascular (0.357) Nervous (0.214) Gastrointestinal (0.143)	Cell Proliferation (0.572) Inflammation (0.214)	pINCY
39	164-208	Nervous (0.280) Reproductive (0.260) Developmental (0.120)	Cell Proliferation (0.740) Inflammation (0.180)	PSPORT1
40	369-411	Cardiovascular (0.250) Developmental (0.250) Gastrointestinal (0.250)	Cell Proliferation (0.500) Inflammation (0.250)	pINCY
41	272-316	Reproductive (0.392) Gastrointestinal (0.118) Hematopoietic/Immune (0.118)	Cell Proliferation (0.626) Inflammation (0.137)	pINCY
42	664-708	Nervous (0.211) Reproductive (0.211) Cardiovascular (0.158)	Cell Proliferation (0.614) Inflammation (0.281)	pINCY
43	226-270	Reproductive (1.000)	Inflammation (1.000)	PBLUESCRIPT
44	11-55	Reproductive (0.254) Gastrointestinal (0.206) Cardiovascular (0.159)	Cell Proliferation (0.698) Inflammation (0.206)	PSPORT1

Table 3 (cont.)

Nucleotide Seq ID NO:	Selected Fragments	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
45	637-681	Reproductive (0.281) Nervous (0.188) Gastrointestinal (0.156)	Cell Proliferation (0.781) Inflammation (0.234)	pINCY
46	1016-1060	Nervous (0.330) Reproductive (0.183) Hematopoietic/Immune (0.122)	Cell Proliferation (0.582) Inflammation (0.235)	pINCY
47	737-781	Nervous (0.218) Reproductive (0.188) Gastrointestinal (0.158)	Cell Proliferation (0.655) Inflammation (0.211)	pINCY
48	469-513	Reproductive (0.222) Hematopoietic/Immune (0.160) Nervous (0.160)	Cell Proliferation (0.543) Inflammation (0.272)	pINCY
49	226-270	Gastrointestinal (0.333) Hematopoietic/Immune (0.333) Reproductive (0.333)	Inflammation (1.000)	pINCY
50	456-500	Reproductive (0.289) Gastrointestinal (0.133) Hematopoietic/Immune (0.133)	Cell Proliferation (0.778) Inflammation (0.156)	PSPORT1
51	252-296	Nervous (0.500) Gastrointestinal (0.200) Cardiovascular (0.100)	Cell Proliferation (1.000) Inflammation (0.200)	PBLUESCRIPT
52	60-104	Nervous (0.326) Reproductive (0.326) Cardiovascular (0.152)	Cell proliferation (0.565) Inflammation (0.369)	pINCY

Table 3 (cont.)

Nucleotide Seq ID NO:	Selected Fragments	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
53	488-532	Reproductive (0.232) Nervous (0.195) Hematopoietic/Immune (0.146)	Cell proliferation (0.622) Inflammation (0.427)	pINCY
54	686-730	Reproductive (0.250) Gastrointestinal (0.150) Hematopoietic/Immune (0.150)	Cell proliferation (0.700) Inflammation (0.400)	pINCY
55	543-587 1299-1343	Reproductive (0.282) Nervous (0.155) Gastrointestinal (0.146)	Cell proliferation (0.592) Inflammation (0.359)	pINCY
56	345-389 792-836	Nervous (0.268) Reproductive (0.169) Cardiovascular (0.113) Hematopoietic/Immune (0.113)	Cell proliferation (0.606) Inflammation (0.296)	pINCY
57	163-207	Reproductive (0.270) Gastrointestinal (0.189) Nervous (0.156)	Cell proliferation (0.705) Inflammation (0.254)	pINCY
58	381-425 726-770	Nervous (0.317) Reproductive (0.250) Gastrointestinal (0.117)	Cell proliferation (0.450) Inflammation (0.283)	pINCY

Table 4

Nucleotide SEQ ID NO:	Library	Library Description
30	SYNORAT04	This library was constructed using RNA isolated from the wrist synovial membrane tissue of a 62-year-old female with rheumatoid arthritis.
31	MENITUT03	This library was constructed using RNA isolated from brain meningioma tissue removed from a 35-year-old female during excision of a cerebral meningeal lesion. Pathology indicated a benign neoplasm in the right cerebellopontine angle of the brain. Patient history included hypothyroidism. Family history included myocardial infarction and breast cancer.
32	PENITUT01	This library was constructed using RNA isolated from tumor tissue removed from the penis of a 64-year-old male during penile amputation. Pathology indicated a fungating invasive grade 4 squamous cell carcinoma involving the inner wall of the foreskin and extending onto the glans penis. Patient history included benign neoplasm of the large bowel, atherosclerotic coronary artery disease, angina pectoris, gout, and obesity. Family history included malignant pharyngeal neoplasm, chronic lymphocytic leukemia, and chronic liver disease.
33	PROSTUT12	This library was constructed using RNA isolated from prostate tumor tissue removed from a 65-year-old male during a radical prostatectomy. Pathology indicated an adenocarcinoma (Gleason grade 2+2). Adenofibromatous hyperplasia was also present. The patient presented with elevated prostate specific antigen (PSA).
34	BONEUNT01	This library was constructed using RNA isolated from Saos-2, a primary osteogenic sarcoma cell line (ATCC HTB-85) derived from an 11-year-old Caucasian female.
35	UTRSNON03	This library was constructed from 6.4 million independent clones from a uterine library. RNA for these libraries was isolated from uterine myometrial tissue removed from a 41-year-old female during a vaginal hysterectomy with dilation and curettage. The endometrium was secretory and contained fragments of endometrial polyps. Benign endo- and ectocervical mucosa were identified in the endocervix. Pathology for the associated tumor tissue indicated uterine leiomyoma. The normalization and hybridization conditions were adapted from Soares et al. (Proc.Natl.Acad.Sci. USA (1994) 91:9928).
36	BRAITUT29	This library was constructed using RNA isolated from brain tumor tissue removed from the parietal lobe of a 43-year-old female during excision of a cerebral meningeal lesion. Pathology indicated high grade glioma. Family history included acute myocardial infarction, atherosclerotic coronary artery disease, benign hypertension, and hyperlipidemia.
37	BMARNOT02	This library was constructed using RNA isolated from the bone marrow of 24 male and female Caucasian donors, 16 to 70 years old. (RNA came from Clontech.)

Table 4 (cont.)

Nucleotide SEQ ID NO:	Library	Library Description
38	THYRNOT03	This library was constructed using RNA isolated from thyroid tissue removed from the left thyroid of a 28-year-old Caucasian female during a complete thyroidectomy. Pathology indicated a small nodule of adenomatous hyperplasia present in the left thyroid. Pathology for the associated tumor tissue indicated dominant follicular adenoma, forming a well-encapsulated mass in the left thyroid.
39	PROSNON01	This normalized library was constructed from 4.4 million independent clones from a prostate library. Starting RNA was made from prostate tissue removed from a 28-yearold Caucasian male who died from a self-inflicted gunshot wound. The normalization and hybridization conditions were adapted from Soares, M.B. et al. (1994) Proc. Natl. Acad. Sci. USA 91:9228-9232, using a longer (19 hour) reannealing hybridization period.
40	LUNGNOT14	This library was constructed using RNA isolated from lung tissue removed from the left lower lobe of a 47-year-old Caucasian male during a segmental lung resection. Pathology for the associated tumor tissue indicated a grade 4 adenocarcinoma, and the parenchyma showed calcified granuloma. Patient history included benign hypertension and chronic obstructive pulmonary disease. Family history included type II diabetes and acute myocardial infarction.
41	PROSTUT12	This library was constructed using RNA isolated from prostate tumor tissue removed from a 65-year-old Caucasian male during a radical prostatectomy. Pathology indicated an adenocarcinoma (Gleason grade 2+2). Adenofibromatous hyperplasia was also present. The patient presented with elevated prostate specific antigen (PSA).
42	LUNGFET03	This library was constructed using RNA isolated from lung tissue removed from a Caucasian female fetus who died at 20 weeks' gestation.
43	TESTNOT03	This library was constructed using RNA isolated from testicular tissue removed from a 37-year-old Caucasian male, who died from liver disease. Patient history included cirrhosis, jaundice, and liver failure.
44	OVARTUT01	This library was constructed using RNA isolated from ovarian tumor tissue removed from a 43-year-old Caucasian female during removal of the fallopian tubes and ovaries. Pathology indicated grade 2 mucinous cystadenocarcinoma involving the entire left ovary. Patient history included mitral valve disorder, pneumonia, and viral hepatitis. Family history included atherosclerotic coronary artery disease, pancreatic cancer, stress reaction, cerebrovascular disease, breast cancer, and uterine cancer.
45	HNT3AZT01	This library was constructed using RNA isolated from the hNT2 cell line (derived from a human teratocarcinoma that exhibited properties characteristic of a committed neuronal precursor). Cells were treated for three days with 0.35 micromolar 5-aza-2'deoxycytidine (AZ).

Table 4 (cont.)

Nucleotide SEQ ID NO:	Library	Library Description
46	PONSAZT01	This library was constructed using RNA isolated from diseased pons tissue from the brain of a 74-year-old Caucasian male who died from Alzheimer's disease.
47	BRSTNOT14	This library was constructed using RNA isolated from breast tissue obtained from a 62-yearold Caucasian female during a unilateral extended simple mastectomy. Pathology for the associated tumor tissue indicated an invasive grade 3 (of 4), nuclear grade 3 (of 3) adenocarcinoma, ductal type. Patient history included a benign colon neoplasm, hyperlipidemia, cardiac dysrhythmia, and obesity. Family history included atherosclerotic coronary artery disease, myocardial infarction, colon cancer, ovarian cancer, lung cancer, and cerebrovascular disease.
48	THYMFET03	This library was constructed using RNA isolated from thymus tissue removed from a Caucasian male fetus.
49	UCMCNOT04	This library was constructed using RNA isolated from mononuclear cells obtained from the umbilical cord blood of multiple individuals of mixed age and sex. The cells were treated with G-CSF.
50	BRSTTUT01	This library was constructed using RNA isolated from breast tumor tissue removed from a 55 year-old Caucasian female during a unilateral extended simple mastectomy. Pathology indicated invasive grade 4 mammary adenocarcinoma of mixed lobular and ductal type, extensively involving the left breast. Family history included benign hypertension, atherosclerotic coronary artery disease, cerebrovascular disease, and depressive disorder.
51	HNT2RAT01	This library was constructed at Stratagene (STR937231), using RNA isolated from the hNT2 cell line (derived from a human teratocarcinoma that exhibited properties characteristic of a committed neuronal precursor). Cells were treated with retinoic acid for 24 hours.
52	MENITUT03	This library was constructed using RNA isolated from brain meningioma tissue removed from a 35-year-old Caucasian female during excision of a cerebral meningeal lesion. Pathology indicated a benign neoplasm in the right cerebellopontine angle of the brain. Patient history included hypothyroidism. Family history included myocardial infarction and breast cancer.
53	SPLNFET02	This library was constructed using RNA isolated from spleen tissue removed from a Caucasian male fetus, who died at 23 weeks' gestation.
54	SEMVNOT01	This library was constructed using RNA isolated from seminal vesicle tissue removed from a 58-year-old Caucasian male during radical prostatectomy. Pathology for the associated tumor tissue indicated adenocarcinoma (Gleason grade 3+2) of the prostate. Adenofibromatous hyperplasia was also present. The patient presented with elevated prostate specific antigen (PSA). Family history included a malignant breast neoplasm.

Table 4 (cont.)

Nucleotide SEQ ID NO:	Library	Library Description
55	CONUTUT01	This library was constructed using RNA isolated from sigmoid mesentery tumor tissue obtained from a 61-year-old female during a total abdominal hysterectomy and bilateral salpingo-oophorectomy with regional lymph node excision. Pathology indicated a metastatic grade 4 malignant mixed mullerian tumor present in the sigmoid mesentery at two sites.
56	HEARFET02	This library was constructed using RNA isolated from heart tissue removed from a Caucasian male fetus, who was stillborn at 23 weeks' gestation with a hypoplastic left heart.
57	BRAIFET01	This library was constructed using RNA isolated from brain tissue removed from a Caucasian male fetus, who was stillborn at 23 weeks' gestation with a hypoplastic left heart.
58	BRAINOT23	This library was constructed using RNA isolated from right temporal lobe tissue removed from a 45-year-old Black male during a brain lobectomy. Pathology for the associated tumor tissue indicated dysembryoplastic neuroepithelial tumor of the right temporal lobe. The right temporal region dura was consistent with calcifying pseudotumor of the neuraxis. The patient presented with convulsive intractable epilepsy, partial epilepsy, and memory disturbance. Patient history included obesity, meningitis, backache, unspecified sleep apnea, acute stress reaction, acquired knee deformity, and chronic sinusitis. Family history included obesity, benign hypertension, cirrhosis of the liver, alcohol abuse, hyperlipidemia, cerebrovascular disease, and type II diabetes.

Table 5

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25: 3389-3402.	*ESTs:* Probability value= 1.0E-8 or less *Full Length sequences*: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises as least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad Sci. 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183: 63-98; and Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489.	*ESTs:* fasta E value=1.06E-6 *Assembled ESTs*: fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less *Full Length sequences:* fastx score=100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S and J.G. Henikoff, Nucl. Acid Res., 19:6565-72, 1991. J.G. Henikoff and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37: 417-424.	Score=1000 or greater; Ratio of Score/Strength = 0.75 or larger; and, if applicable, Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol., 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322.	Score=10-50 bits for PFAM hits, depending on individual protein families

Table 5 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25: 217-221.	Normalized quality score≥GCG-specified "HIGH" value for that particular Prosite motif. Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M. S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12: 431-439.	Score=3.5 or greater
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch et al. <u>supra</u>; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

What is claimed is:

1. A substantially purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-29 and fragments thereof.

2. A substantially purified variant having at least 90% amino acid sequence identity to the amino acid sequence of claim 1.

3. An isolated and purified polynucleotide encoding the polypeptide of claim 1.

4. An isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide of claim 3.

5. An isolated and purified polynucleotide which hybridizes under stringent conditions to the polynucleotide of claim 3.

6. An isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide of claim 3.

7. A method for detecting a polynucleotide, the method comprising the steps of:

(a) hybridizing the polynucleotide of claim 6 to at least one nucleic acid in a sample, thereby forming a hybridization complex; and

(b) detecting the hybridization complex, wherein the presence of the hybridization complex correlates with the presence of the polynucleotide in the sample.

8. The method of claim 7 further comprising amplifying the polynucleotide prior to hybridization.

9. An isolated and purified polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:30-58 and fragments thereof.

10. An isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide of claim 9.

11. An isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide of claim 9.

12. An expression vector comprising at least a fragment of the polynucleotide of claim 3.

13. A host cell comprising the expression vector of claim 12.

14. A method for producing a polypeptide, the method comprising the steps of:

a) culturing the host cell of claim 13 under conditions suitable for the expression of the polypeptide; and

b) recovering the polypeptide from the host cell culture.

15. A pharmaceutical composition comprising the polypeptide of claim 1 in conjunction with a suitable pharmaceutical carrier.

16. A purified antibody which specifically binds to the polypeptide of claim 1.

17. A purified agonist of the polypeptide of claim 1.

18. A purified antagonist of the polypeptide of claim 1.

19. A method for treating or preventing a disorder associated with decreased expression or activity of GTPAP, the method comprising administering to a subject in need of such treatment an effective amount of the pharmaceutical composition of claim 15.

20. A method for treating or preventing a disorder associated with increased expression or activity of GTPAP, the method comprising administering to a subject in need of such treatment an effective amount of the antagonist of claim 18.

SEQUENCE LISTING

```
<110> INCYTE PHARMACEUTICALS, INC.
      HILLMAN, Jennifer L.
      TANG, Y. Tom
      BANDMAN, Olga
      LAL, Preeti
      YUE, Henry
      LU, Dyung Aina M.
      BAUGHN, Mariah R.
      YANG, Junming
      AZIMZAI, Yalda

<120> GTPASE ASSOCIATED PROTEINS

<130> PF-0629 PCT

<140> To Be Assigned
<141> Herewith

<150> 60/109,592; 60/118,610; 60/127,990
<151> 1998-11-23; 1999-02-04; 1999-04-06

<160> 58

<170> PERL Program

<210> 1
<211> 1002
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 708398CD1

<400> 1
Met Pro Ser Lys Phe Ser Cys Arg Gln Leu Arg Glu Ala Gly Gln
  1               5                  10                  15
Cys Phe Glu Ser Phe Leu Val Val Arg Gly Leu Asp Met Glu Thr
                 20                  25                  30
Asp Arg Glu Arg Leu Arg Thr Ile Tyr Asn Arg Asp Phe Lys Ile
                 35                  40                  45
Ser Phe Gly Thr Pro Ala Pro Gly Phe Ser Ser Met Leu Tyr Gly
                 50                  55                  60
Met Lys Ile Ala Asn Leu Ala Tyr Val Thr Lys Thr Arg Val Arg
                 65                  70                  75
Phe Phe Arg Leu Asp Arg Trp Ala Asp Val Arg Phe Pro Glu Lys
                 80                  85                  90
Arg Arg Met Lys Leu Gly Ser Asp Ile Ser Lys His His Lys Ser
                 95                 100                 105
Leu Leu Ala Lys Ile Phe Tyr Asp Arg Ala Glu Tyr Leu His Gly
                110                 115                 120
Lys His Gly Val Asp Val Glu Val Gln Gly Pro His Glu Ala Arg
```

```
                125                 130                 135
Asp Gly Gln Leu Leu Ile Arg Leu Asp Leu Asn Arg Lys Glu Val
                140                 145                 150
Leu Thr Leu Arg Leu Arg Asn Gly Gly Thr Gln Ser Val Thr Leu
                155                 160                 165
Thr His Leu Phe Pro Leu Cys Arg Thr Pro Gln Phe Ala Phe Tyr
                170                 175                 180
Asn Glu Asp Gln Glu Leu Pro Cys Pro Leu Gly Pro Gly Glu Cys
                185                 190                 195
Tyr Glu Leu His Val His Cys Lys Thr Ser Phe Val Gly Tyr Phe
                200                 205                 210
Pro Ala Thr Val Leu Trp Glu Leu Leu Gly Pro Gly Glu Ser Gly
                215                 220                 225
Ser Glu Gly Ala Gly Thr Phe Tyr Ile Ala Arg Phe Leu Ala Ala
                230                 235                 240
Val Ala His Ser Pro Leu Ala Ala Gln Leu Lys Pro Met Thr Pro
                245                 250                 255
Phe Lys Arg Thr Arg Ile Thr Gly Asn Pro Val Val Thr Asn Arg
                260                 265                 270
Ile Glu Glu Gly Glu Arg Pro Asp Arg Ala Lys Gly Tyr Asp Leu
                275                 280                 285
Glu Leu Ser Met Ala Leu Gly Thr Tyr Tyr Pro Pro Pro Arg Leu
                290                 295                 300
Arg Gln Leu Leu Pro Met Leu Leu Gln Gly Thr Ser Ile Phe Thr
                305                 310                 315
Ala Pro Lys Glu Ile Ala Glu Ile Lys Ala Gln Leu Glu Thr Ala
                320                 325                 330
Leu Lys Trp Arg Asn Tyr Glu Val Lys Leu Arg Leu Leu Leu His
                335                 340                 345
Leu Glu Glu Leu Gln Met Glu His Asp Ile Arg His Tyr Asp Leu
                350                 355                 360
Glu Ser Val Pro Met Thr Trp Asp Pro Val Asp Gln Asn Pro Arg
                365                 370                 375
Leu Leu Thr Leu Glu Val Pro Gly Val Thr Glu Ser Arg Pro Ser
                380                 385                 390
Val Leu Arg Gly Asp His Leu Phe Ala Leu Leu Ser Ser Glu Thr
                395                 400                 405
His Gln Glu Asp Pro Ile Thr Tyr Lys Gly Phe Val His Lys Val
                410                 415                 420
Glu Leu Asp Arg Val Lys Leu Ser Phe Ser Met Ser Leu Leu Ser
                425                 430                 435
Arg Phe Val Asp Gly Leu Thr Phe Lys Val Asn Phe Thr Phe Asn
                440                 445                 450
Arg Gln Pro Leu Arg Val Gln His Arg Ala Leu Glu Leu Thr Gly
                455                 460                 465
Arg Trp Leu Leu Trp Pro Met Leu Phe Pro Val Ala Pro Arg Asp
                470                 475                 480
Val Pro Leu Leu Pro Ser Asp Val Lys Leu Lys Leu Tyr Asp Arg
                485                 490                 495
Ser Leu Glu Ser Asn Pro Glu Gln Leu Gln Ala Met Arg His Ile
                500                 505                 510
Val Thr Gly Thr Thr Arg Pro Ala Pro Tyr Ile Ile Phe Gly Pro
                515                 520                 525
Pro Gly Thr Gly Lys Thr Val Thr Leu Val Glu Ala Ile Lys Gln
                530                 535                 540
```

Val Val Lys His Leu Pro Lys Ala His Ile Leu Ala Cys Ala Pro
545 550 555
Ser Asn Ser Gly Ala Asp Leu Leu Cys Gln Arg Leu Arg Val His
560 565 570
Leu Pro Ser Ser Ile Tyr Arg Leu Leu Ala Pro Ser Arg Asp Ile
575 580 585
Arg Met Val Pro Glu Asp Ile Lys Pro Cys Cys Asn Trp Asp Ala
590 595 600
Lys Lys Gly Glu Tyr Val Phe Pro Ala Lys Lys Lys Leu Gln Glu
605 610 615
Tyr Arg Val Leu Ile Thr Thr Leu Ile Thr Ala Gly Arg Leu Val
620 625 630
Ser Ala Gln Phe Pro Ile Asp His Phe Thr His Ile Phe Ile Asp
635 640 645
Glu Ala Gly His Cys Met Glu Pro Glu Ser Leu Val Ala Ile Ala
650 655 660
Gly Leu Met Glu Val Lys Glu Thr Gly Asp Pro Gly Gly Gln Leu
665 670 675
Val Leu Ala Gly Asp Pro Arg Gln Leu Gly Pro Val Leu Arg Ser
680 685 690
Pro Leu Thr Gln Lys His Gly Leu Gly Tyr Ser Leu Leu Glu Arg
695 700 705
Leu Leu Ile Tyr Asn Ser Leu Tyr Lys Lys Gly Pro Asp Gly Tyr
710 715 720
Asp Pro Gln Phe Ile Thr Lys Leu Leu Arg Asn Tyr Arg Ser His
725 730 735
Pro Thr Ile Leu Asp Ile Pro Asn Gln Leu Tyr Tyr Glu Gly Glu
740 745 750
Leu Gln Ala Cys Ala Asp Val Val Asp Arg Glu Arg Phe Cys Arg
755 760 765
Trp Ala Gly Leu Pro Arg Gln Gly Phe Pro Ile Ile Phe His Gly
770 775 780
Val Met Gly Lys Asp Glu Arg Glu Gly Asn Ser Pro Ser Phe Phe
785 790 795
Asn Pro Glu Glu Ala Ala Thr Val Thr Ser Tyr Leu Lys Leu Leu
800 805 810
Leu Ala Pro Ser Ser Lys Lys Gly Lys Ala Arg Leu Ser Pro Arg
815 820 825
Ser Val Gly Val Ile Ser Pro Tyr Arg Lys Gln Val Glu Lys Ile
830 835 840
Arg Tyr Cys Ile Thr Lys Leu Asp Arg Glu Leu Arg Gly Leu Asp
845 850 855
Asp Ile Lys Asp Leu Lys Val Gly Ser Val Glu Glu Phe Gln Gly
860 865 870
Gln Glu Arg Ser Val Ile Leu Ile Ser Thr Val Arg Ser Ser Gln
875 880 885
Ser Phe Val Gln Leu Asp Leu Asp Phe Asn Leu Gly Phe Leu Lys
890 895 900
Asn Pro Lys Arg Phe Asn Val Ala Val Thr Arg Ala Lys Ala Leu
905 910 915
Leu Ile Ile Val Gly Asn Pro Leu Leu Leu Gly His Asp Pro Asp
920 925 930
Trp Lys Val Phe Leu Glu Phe Cys Lys Glu Asn Gly Gly Tyr Thr
935 940 945
Gly Cys Pro Phe Pro Ala Lys Leu Asp Leu Gln Gln Gly Gln Asn

950 955 960
Leu Leu Gln Gly Leu Ser Lys Leu Ser Pro Ser Thr Ser Gly Pro
965 970 975
His Ser His Asp Tyr Leu Pro Gln Glu Arg Glu Gly Glu Gly Gly
980 985 990
Leu Ser Leu Gln Val Glu Pro Glu Trp Arg Asn Glu
995 1000

<210> 2
<211> 338
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1259937CD1

<400> 2
Met Ala Ala Leu Ala Gln Glu Asp Gly Trp Thr Lys Gly Gln Val
1 5 10 15
Leu Val Lys Val Asn Ser Ala Gly Asp Ala Ile Gly Leu Gln Pro
20 25 30
Asp Ala Arg Gly Val Ala Thr Ser Leu Gly Leu Asn Glu Arg Leu
35 40 45
Phe Val Val Asn Pro Gln Glu Val His Glu Leu Ile Pro His Pro
50 55 60
Asp Gln Leu Gly Pro Thr Val Gly Ser Ala Glu Gly Leu Asp Leu
65 70 75
Val Ser Ala Lys Asp Leu Ala Gly Gln Leu Thr Asp His Asp Trp
80 85 90
Ser Leu Phe Asn Ser Ile His Gln Val Glu Leu Ile His Tyr Val
95 100 105
Leu Gly Pro Gln His Leu Arg Asp Val Thr Thr Ala Asn Leu Glu
110 115 120
Arg Phe Met Arg Arg Phe Asn Glu Leu Gln Tyr Trp Val Ala Thr
125 130 135
Glu Leu Cys Leu Cys Pro Val Pro Gly Pro Arg Ala Gln Leu Leu
140 145 150
Arg Lys Phe Ile Lys Leu Ala Ala His Leu Lys Glu Gln Lys Asn
155 160 165
Leu Asn Ser Phe Phe Ala Val Met Phe Gly Leu Ser Asn Ser Ala
170 175 180
Ile Ser Arg Leu Ala His Thr Trp Glu Arg Leu Pro His Lys Val
185 190 195
Arg Lys Leu Tyr Ser Ala Leu Glu Arg Leu Leu Asp Pro Ser Trp
200 205 210
Asn His Arg Val Tyr Arg Leu Ala Leu Ala Lys Leu Ser Pro Pro
215 220 225
Val Ile Pro Phe Met Pro Leu Leu Leu Lys Asp Met Thr Phe Ile
230 235 240
His Glu Gly Asn His Thr Leu Val Glu Asn Leu Ile Asn Phe Glu
245 250 255
Lys Met Arg Met Met Ala Arg Ala Ala Arg Met Leu His His Cys
260 265 270

```
Arg Ser His Asn Pro Val Pro Leu Ser Pro Leu Arg Ser Arg Val
                275                 280                 285
Ser His Leu His Glu Asp Ser Gln Val Ala Arg Ile Ser Thr Cys
                290                 295                 300
Ser Glu Gln Ser Leu Ser Thr Arg Ser Pro Ala Ser Thr Trp Ala
                305                 310                 315
Tyr Val Gln Gln Leu Lys Val Ile Asp Asn Gln Arg Glu Leu Ser
                320                 325                 330
Arg Leu Ser Arg Glu Leu Glu Pro
                335


<210> 3
<211> 211
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No:  1452285CD1

<400> 3
Met Gln Ala Pro His Lys Glu His Leu Tyr Lys Leu Leu Val Ile
  1               5                  10                  15
Gly Asp Leu Gly Val Gly Lys Thr Ser Ile Ile Lys Arg Tyr Val
                 20                  25                  30
His Gln Asn Phe Ser Ser His Tyr Arg Ala Thr Ile Gly Val Asp
                 35                  40                  45
Phe Ala Leu Lys Val Leu His Trp Asp Pro Glu Thr Val Val Arg
                 50                  55                  60
Leu Gln Leu Trp Asp Ile Ala Gly Gln Glu Arg Phe Gly Asn Met
                 65                  70                  75
Thr Arg Val Tyr Tyr Arg Glu Ala Met Gly Ala Phe Ile Val Phe
                 80                  85                  90
Asp Val Thr Arg Pro Ala Thr Phe Glu Ala Val Ala Lys Trp Lys
                 95                 100                 105
Asn Asp Leu Asp Ser Lys Leu Ser Leu Pro Asn Gly Lys Pro Val
                110                 115                 120
Ser Val Val Leu Leu Ala Asn Lys Cys Asp Gln Gly Lys Asp Val
                125                 130                 135
Leu Met Asn Asn Gly Leu Lys Met Asp Gln Phe Cys Lys Glu His
                140                 145                 150
Gly Phe Val Gly Trp Phe Glu Thr Ser Ala Lys Glu Asn Ile Asn
                155                 160                 165
Ile Asp Glu Ala Ser Arg Cys Leu Val Lys His Ile Leu Ala Asn
                170                 175                 180
Glu Cys Asp Leu Met Glu Ser Ile Glu Pro Asp Val Val Lys Pro
                185                 190                 195
His Leu Thr Ser Thr Lys Val Ala Ser Cys Ser Gly Cys Ala Lys
                200                 205                 210
Ser


<210> 4
<211> 516
```

<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1812894CD1

<400> 4
```
Met Glu Thr Met Lys Ala Val Ala Glu Val Ser Glu Ser Thr Lys
 1               5                  10                  15
Ala Glu Ala Val Ala Ala Val Gln Arg Gln Cys Gln Glu Glu Val
                20                  25                  30
Ala Ser Leu Gln Ala Ile Leu Lys Asp Ser Ile Ser Ser Tyr Glu
                35                  40                  45
Ala Gln Ile Thr Ala Leu Lys Gln Glu Arg Gln Gln Gln Gln Gln
                50                  55                  60
Asp Cys Glu Glu Lys Glu Arg Glu Leu Gly Arg Leu Lys Gln Leu
                65                  70                  75
Leu Ser Arg Ala Tyr Pro Leu Asp Ser Leu Glu Lys Gln Met Glu
                80                  85                  90
Lys Ala His Glu Asp Ser Glu Lys Leu Arg Glu Ile Val Leu Pro
                95                 100                 105
Met Glu Lys Glu Ile Glu Glu Leu Lys Ala Lys Leu Leu Arg Ala
               110                 115                 120
Glu Glu Leu Ile Gln Glu Ile Gln Arg Arg Pro Arg His Ala Pro
               125                 130                 135
Ser Leu His Gly Ser Thr Glu Leu Leu Pro Leu Ser Arg Asp Pro
               140                 145                 150
Ser Pro Pro Leu Glu Pro Leu Glu Glu Leu Ser Gly Asp Gly Gly
               155                 160                 165
Pro Ala Ala Glu Ala Phe Ala His Asn Cys Asp Asp Ser Ala Ser
               170                 175                 180
Ile Ser Ser Phe Ser Leu Gly Gly Gly Val Gly Ser Ser Ser Ser
               185                 190                 195
Leu Pro Gln Ser Arg Gln Gly Leu Ser Pro Glu Gln Glu Glu Thr
               200                 205                 210
Ala Ser Leu Val Ser Thr Gly Thr Leu Val Pro Glu Gly Ile Tyr
               215                 220                 225
Leu Pro Pro Pro Gly Tyr Gln Leu Val Pro Asp Thr Gln Trp Glu
               230                 235                 240
Gln Leu Gln Thr Glu Gly Arg Gln Leu Gln Lys Asp Leu Glu Ser
               245                 250                 255
Val Ser Arg Glu Arg Asp Glu Leu Gln Glu Gly Leu Arg Arg Ser
               260                 265                 270
Asn Glu Asp Cys Ala Lys Gln Met Gln Val Leu Leu Ala Gln Val
               275                 280                 285
Gln Asn Ser Glu Gln Leu Leu Arg Thr Leu Gln Gly Thr Val Ser
               290                 295                 300
Gln Ala Gln Glu Arg Val Gln Leu Gln Met Ala Glu Leu Val Thr
               305                 310                 315
Thr His Lys Cys Leu His His Glu Val Lys Arg Leu Asn Glu Glu
               320                 325                 330
Asn Gln Gly Leu Arg Ala Glu Gln Leu Pro Ser Ser Ala Pro Gln
               335                 340                 345
Gly Ser Gln Gln Glu Gln Gly Glu Glu Glu Ser Leu Pro Ser Ser
```

```
                350                 355                 360
Val Pro Glu Leu Gln Gln Leu Leu Cys Cys Thr Arg Gln Glu Ala
                365                 370                 375
Arg Ala Arg Leu Gln Ala Gln Glu His Gly Ala Glu Arg Leu Arg
                380                 385                 390
Ile Glu Ile Val Thr Leu Arg Glu Ala Leu Glu Glu Glu Thr Val
                395                 400                 405
Ala Arg Ala Ser Leu Glu Gly Gln Leu Arg Val Gln Arg Glu Glu
                410                 415                 420
Thr Glu Val Leu Glu Ala Ser Leu Cys Ser Leu Arg Thr Glu Met
                425                 430                 435
Glu Arg Val Gln Gln Glu Gln Ser Lys Ala Gln Leu Pro Asp Leu
                440                 445                 450
Leu Ser Glu Gln Arg Ala Lys Val Leu Arg Leu Gln Ala Glu Leu
                455                 460                 465
Glu Thr Ser Glu Gln Val Gln Arg Asp Phe Val Arg Leu Ser Gln
                470                 475                 480
Ala Leu Gln Val Arg Leu Glu Arg Ile Arg Gln Ala Glu Thr Leu
                485                 490                 495
Glu Gln Val Arg Ser Ile Met Asp Glu Ala Pro Leu Thr Asp Val
                500                 505                 510
Arg Asp Ile Lys Asp Thr
                515


<210> 5
<211> 445
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No:  3074884CD1

<400> 5
Met Pro Glu Asp Ala Asp Glu Asn Ala Glu Glu Glu Leu Leu Arg
  1               5                  10                  15
Gly Glu Pro Leu Leu Pro Ala Gly Thr Gln Arg Val Cys Leu Val
                 20                  25                  30
His Pro Asp Val Lys Trp Gly Pro Gly Lys Ser Gln Met Thr Arg
                 35                  40                  45
Ala Glu Trp Gln Val Ala Glu Ala Thr Ala Leu Val His Thr Leu
                 50                  55                  60
Asp Gly Trp Ser Val Val Gln Thr Met Val Val Ser Thr Lys Thr
                 65                  70                  75
Pro Asp Arg Lys Leu Ile Phe Gly Lys Gly Asn Phe Glu His Leu
                 80                  85                  90
Thr Glu Lys Ile Arg Gly Ser Pro Asp Val Thr Cys Val Phe Leu
                 95                 100                 105
Asn Val Glu Arg Met Ala Ala Pro Thr Lys Lys Glu Leu Glu Ala
                110                 115                 120
Ala Trp Gly Val Glu Val Phe Asp Arg Phe Thr Val Val Leu His
                125                 130                 135
Ile Phe Arg Cys Asn Ala Arg Thr Lys Glu Ala Arg Leu Gln Val
                140                 145                 150
```

```
Ala Leu Ala Glu Met Pro Leu His Arg Ser Asn Leu Lys Arg Asp
                155                 160                 165
Val Ala His Leu Tyr Arg Gly Val Gly Ser Arg Tyr Ile Met Gly
                170                 175                 180
Ser Gly Glu Ser Phe Met Gln Leu Gln Gln Arg Leu Leu Arg Glu
                185                 190                 195
Lys Glu Ala Lys Ile Arg Lys Ala Leu Asp Arg Leu Arg Lys Lys
                200                 205                 210
Arg His Leu Leu Arg Arg Gln Arg Thr Arg Arg Glu Phe Pro Val
                215                 220                 225
Ile Ser Val Val Gly Tyr Thr Asn Cys Gly Lys Thr Thr Leu Ile
                230                 235                 240
Lys Ala Leu Thr Gly Asp Ala Ala Ile Gln Pro Arg Asp Gln Leu
                245                 250                 255
Phe Ala Thr Leu Asp Val Thr Ala His Ala Gly Thr Leu Pro Ser
                260                 265                 270
Arg Met Thr Val Leu Tyr Val Asp Thr Ile Gly Phe Leu Ser Gln
                275                 280                 285
Leu Pro His Gly Leu Ile Glu Ser Phe Ser Ala Thr Leu Glu Asp
                290                 295                 300
Val Ala His Ser Asp Leu Ile Leu His Val Arg Asp Val Ser His
                305                 310                 315
Pro Glu Ala Glu Leu Gln Lys Cys Ser Val Leu Ser Thr Leu Arg
                320                 325                 330
Gly Leu Gln Leu Pro Ala Pro Leu Leu Asp Ser Met Val Glu Val
                335                 340                 345
His Asn Lys Val Asp Leu Val Pro Gly Tyr Ser Pro Thr Glu Pro
                350                 355                 360
Asn Val Val Pro Val Ser Ala Leu Arg Gly His Gly Leu Gln Glu
                365                 370                 375
Leu Lys Ala Glu Leu Asp Ala Ala Val Leu Lys Ala Thr Gly Arg
                380                 385                 390
Gln Ile Leu Thr Leu Arg Val Arg Leu Ala Gly Ala Gln Leu Ser
                395                 400                 405
Trp Leu Tyr Lys Glu Ala Thr Val Gln Glu Val Asp Val Ile Pro
                410                 415                 420
Glu Asp Gly Ala Ala Asp Val Arg Val Ile Ile Ser Asn Ser Ala
                425                 430                 435
Tyr Gly Lys Phe Arg Lys Leu Phe Pro Gly
                440                 445
```

<210> 6
<211> 445
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3452277CD1

<400> 6

```
Met Tyr Tyr Gln Gln Ala Leu Met Arg Ser Thr Val Lys Ser Ser
  1               5                  10                  15
Val Ser Leu Gly Gly Ile Val Lys Tyr Ser Glu Gln Phe Ser Ser
```

```
                20                  25                  30
Asn Asp Ala Ile Met Ser Gly Cys Leu Pro Ser Asn Pro Trp Ile
                35                  40                  45
Thr Asp Asp Thr Gln Phe Trp Asp Leu Asn Ala Lys Leu Val Glu
                50                  55                  60
Ile Pro Thr Lys Met Arg Val Glu Arg Trp Ala Phe Asn Phe Ser
                65                  70                  75
Glu Leu Ile Arg Asp Pro Lys Gly Arg Gln Ser Phe Gln Tyr Phe
                80                  85                  90
Leu Lys Lys Glu Phe Ser Gly Glu Asn Leu Gly Phe Trp Glu Ala
                95                  100                 105
Cys Glu Asp Leu Lys Tyr Gly Asp Gln Ser Lys Val Lys Glu Lys
                110                 115                 120
Ala Glu Glu Ile Tyr Lys Leu Phe Leu Ala Pro Gly Ala Arg Arg
                125                 130                 135
Trp Ile Asn Ile Asp Gly Lys Thr Met Asp Ile Thr Val Lys Gly
                140                 145                 150
Leu Lys His Pro His Arg Tyr Val Leu Asp Ala Ala Gln Thr His
                155                 160                 165
Ile Tyr Met Leu Met Lys Lys Asp Ser Tyr Ala Arg Tyr Leu Lys
                170                 175                 180
Ser Pro Ile Tyr Lys Asp Met Leu Ala Lys Ala Ile Glu Pro Gln
                185                 190                 195
Glu Thr Thr Lys Lys Ser Ser Thr Leu Pro Phe Met Arg Arg His
                200                 205                 210
Leu Arg Ser Ser Pro Ser Pro Val Ile Leu Arg Gln Leu Glu Glu
                215                 220                 225
Glu Ala Lys Ala Arg Glu Ala Ala Asn Thr Val Asp Ile Thr Gln
                230                 235                 240
Pro Gly Gln His Met Ala Pro Ser Pro His Leu Thr Val Tyr Thr
                245                 250                 255
Gly Thr Cys Met Pro Pro Ser Pro Ser Ser Pro Phe Ser Ser Ser
                260                 265                 270
Cys Arg Ser Pro Arg Lys Pro Phe Ala Ser Pro Ser Arg Phe Ile
                275                 280                 285
Arg Arg Pro Ser Thr Thr Ile Cys Pro Ser Pro Ile Arg Val Ala
                290                 295                 300
Leu Glu Ser Ser Ser Gly Leu Glu Gln Lys Gly Glu Cys Ser Gly
                305                 310                 315
Ser Met Ala Pro Arg Gly Pro Ser Val Thr Glu Ser Ser Glu Ala
                320                 325                 330
Ser Leu Asp Thr Ser Trp Pro Arg Ser Arg Pro Arg Ala Pro Pro
                335                 340                 345
Lys Ala Arg Met Ala Leu Ser Phe Ser Arg Phe Leu Arg Arg Gly
                350                 355                 360
Cys Leu Ala Ser Pro Val Phe Ala Arg Leu Ser Pro Lys Cys Pro
                365                 370                 375
Ala Val Ser His Gly Arg Val Gln Pro Leu Gly Asp Val Gly Gln
                380                 385                 390
Gln Leu Pro Arg Leu Lys Ser Lys Arg Val Ala Asn Phe Phe Gln
                395                 400                 405
Ile Lys Met Asp Val Pro Thr Gly Ser Gly Thr Cys Leu Met Asp
                410                 415                 420
Ser Glu Asp Ala Gly Thr Gly Glu Ser Gly Asp Arg Ala Thr Glu
                425                 430                 435
```

Lys Glu Val Ile Cys Pro Trp Glu Ser Leu
440 445

<210> 7
<211> 281
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4203832CD1

<400> 7
Met Lys Leu Ala Ala Met Ile Lys Lys Met Cys Pro Ser Asp Ser
1 5 10 15
Glu Leu Ser Ile Pro Ala Lys Asn Cys Tyr Arg Met Val Ile Leu
20 25 30
Gly Ser Ser Lys Val Gly Lys Thr Ala Ile Val Ser Arg Phe Leu
35 40 45
Thr Gly Arg Phe Glu Asp Ala Tyr Thr Pro Thr Ile Glu Asp Phe
50 55 60
His Arg Lys Phe Tyr Ser Ile Arg Gly Glu Val Tyr Gln Leu Asp
65 70 75
Ile Leu Asp Thr Ser Gly Asn His Pro Phe Pro Ala Met Arg Cys
80 85 90
Leu Ser Ile Leu Thr Gly Asp Val Phe Ile Leu Val Phe Ser Leu
95 100 105
Asp Asn Arg Asp Ser Phe Glu Glu Val Gln Arg Leu Arg Gln Gln
110 115 120
Ile Leu Asp Thr Lys Ser Cys Leu Lys Asn Lys Thr Lys Glu Asn
125 130 135
Val Asp Val Pro Leu Val Ile Cys Gly Asn Lys Gly Asp Arg Asp
140 145 150
Phe Tyr Arg Glu Val Asp Gln Arg Glu Ile Glu Gln Leu Val Gly
155 160 165
Asp Asp Pro Gln Arg Cys Ala Tyr Phe Glu Ile Ser Ala Lys Lys
170 175 180
Asn Ser Ser Leu Asp Gln Met Phe Arg Ala Leu Phe Ala Met Ala
185 190 195
Lys Leu Pro Ser Glu Met Ser Pro Asp Leu His Arg Lys Val Ser
200 205 210
Val Gln Tyr Cys Asp Val Leu His Lys Lys Ala Leu Arg Asn Lys
215 220 225
Lys Leu Leu Arg Ala Gly Ser Gly Gly Gly Gly Gly Asp Pro Gly
230 235 240
Asp Ala Phe Gly Ile Val Ala Pro Phe Ala Arg Arg Pro Ser Val
245 250 255
His Ser Asp Leu Met Tyr Ile Arg Glu Lys Ala Ser Ala Gly Ser
260 265 270
Gln Ala Lys Asp Lys Glu Arg Cys Val Ile Ser
275 280
<210> 8
<211> 301
<212> PRT

<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 104368CD1

<400> 8
Met Thr Thr Leu Asp Asp Lys Leu Leu Gly Glu Lys Leu Gln Tyr
1 5 10 15
Tyr Tyr Ser Ser Ser Glu Asp Glu Asp Ser Asp His Glu Asp Lys
20 25 30
Asp Arg Gly Arg Cys Ala Pro Ala Ser Ser Ser Val Pro Ala Glu
35 40 45
Ala Glu Leu Ala Gly Glu Gly Ile Ser Val Asn Thr Gly Pro Lys
50 55 60
Gly Val Ile Asn Asp Trp Arg Arg Phe Lys Gln Leu Glu Thr Glu
65 70 75
Gln Arg Glu Glu Gln Cys Arg Glu Met Glu Arg Leu Ile Lys Lys
80 85 90
Leu Ser Met Thr Cys Arg Ser His Leu Asp Glu Glu Glu Glu Gln
95 100 105
Gln Lys Gln Lys Asp Leu Gln Glu Lys Ile Ser Gly Lys Met Thr
110 115 120
Leu Lys Glu Phe Ala Ile Met Asn Glu Asp Gln Asp Asp Glu Glu
125 130 135
Phe Leu Gln Gln Tyr Arg Lys Gln Arg Met Glu Glu Met Arg Gln
140 145 150
Gln Leu His Lys Gly Pro Gln Phe Lys Gln Val Phe Glu Ile Ser
155 160 165
Ser Gly Glu Gly Phe Leu Asp Met Ile Asp Lys Glu Gln Lys Ser
170 175 180
Ile Val Ile Met Val His Ile Tyr Glu Asp Gly Ile Pro Gly Thr
185 190 195
Glu Ala Met Asn Gly Cys Met Ile Cys Leu Ala Ala Glu Tyr Pro
200 205 210
Ala Val Lys Phe Cys Lys Val Lys Ser Ser Val Ile Gly Ala Ser
215 220 225
Ser Gln Phe Thr Arg Asn Ala Leu Pro Ala Leu Leu Ile Tyr Lys
230 235 240
Gly Gly Glu Leu Ile Gly Asn Phe Val Arg Val Thr Asp Gln Leu
245 250 255
Gly Asp Asp Phe Phe Ala Val Asp Leu Glu Ala Phe Leu Gln Glu
260 265 270
Phe Gly Leu Leu Pro Glu Lys Glu Val Leu Val Leu Thr Ser Val
275 280 285
Arg Asn Ser Ala Thr Cys His Ser Glu Asp Ser Asp Leu Glu Ile
290 295 300
Asp

<210> 9
<211> 485
<212> PRT

```
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No:  1441680CD1

<400> 9
Met Arg Ala Val Leu Thr Trp Arg Asp Lys Ala Glu His Cys Ile
 1               5                  10                  15
Asn Asp Ile Ala Phe Lys Pro Asp Gly Thr Gln Leu Ile Leu Ala
                20                  25                  30
Ala Gly Ser Arg Leu Leu Val Tyr Asp Thr Ser Asp Gly Thr Leu
                35                  40                  45
Leu Gln Pro Leu Lys Gly His Lys Asp Thr Val Tyr Cys Val Ala
                50                  55                  60
Tyr Ala Lys Asp Gly Lys Arg Phe Ala Ser Gly Ser Ala Asp Lys
                65                  70                  75
Ser Val Ile Ile Trp Thr Ser Lys Leu Glu Gly Ile Leu Lys Tyr
                80                  85                  90
Thr His Asn Asp Ala Ile Gln Cys Val Ser Tyr Asn Pro Ile Thr
                95                 100                 105
His Gln Leu Ala Ser Cys Ser Ser Ser Asp Phe Gly Leu Trp Ser
               110                 115                 120
Pro Glu Gln Lys Ser Val Ser Lys His Lys Ser Ser Ser Lys Ile
               125                 130                 135
Ile Cys Cys Ser Trp Thr Asn Asp Gly Gln Tyr Leu Ala Leu Gly
               140                 145                 150
Met Phe Asn Gly Ile Ile Ser Ile Arg Asn Lys Asn Gly Glu Glu
               155                 160                 165
Lys Val Lys Ile Glu Arg Pro Gly Gly Ser Leu Ser Pro Ile Trp
               170                 175                 180
Ser Ile Cys Trp Asn Pro Ser Arg Glu Glu Arg Asn Asp Ile Leu
               185                 190                 195
Ala Val Ala Asp Trp Gly Gln Lys Val Ser Phe Tyr Gln Leu Ser
               200                 205                 210
Gly Lys Gln Ile Gly Lys Asp Arg Ala Leu Asn Phe Asp Pro Cys
               215                 220                 225
Cys Ile Ser Tyr Phe Thr Lys Gly Glu Tyr Ile Leu Leu Gly Gly
               230                 235                 240
Ser Asp Lys Gln Val Ser Leu Phe Thr Lys Asp Gly Val Arg Leu
               245                 250                 255
Gly Thr Val Gly Glu Gln Asn Ser Trp Val Trp Thr Cys Gln Ala
               260                 265                 270
Lys Pro Asp Ser Asn Tyr Val Val Val Gly Cys Gln Asp Gly Thr
               275                 280                 285
Ile Ser Phe Tyr Gln Leu Ile Phe Ser Thr Val His Gly Val Tyr
               290                 295                 300
Lys Asp Arg Tyr Ala Tyr Arg Asp Ser Met Thr Asp Val Ile Val
               305                 310                 315
Gln His Leu Ile Thr Glu Gln Lys Val Arg Ile Lys Cys Lys Glu
               320                 325                 330
Leu Val Lys Lys Ile Ala Ile Tyr Arg Asn Arg Leu Ala Ile Gln
               335                 340                 345
Leu Pro Glu Lys Ile Leu Ile Tyr Glu Leu Tyr Ser Glu Asp Leu
               350                 355                 360
```

```
Ser Asp Met His Tyr Arg Val Lys Glu Lys Ile Ile Lys Lys Phe
                365                 370                 375
Glu Cys Asn Leu Leu Val Val Cys Ala Asn His Ile Ile Leu Cys
                380                 385                 390
Gln Glu Lys Arg Leu Gln Cys Leu Ser Phe Ser Gly Val Lys Glu
                395                 400                 405
Arg Glu Trp Gln Met Glu Ser Leu Ile Arg Tyr Ile Lys Val Ile
                410                 415                 420
Gly Gly Pro Pro Gly Arg Glu Gly Leu Leu Val Gly Leu Lys Lys
                425                 430                 435
Met Tyr Leu Leu Val Tyr Ser Phe Ile Leu Ile Val Lys Asp Tyr
                440                 445                 450
Phe Ser Leu Ser Thr Asp Val Leu Gly Asn Leu Thr Trp Lys His
                455                 460                 465
Val Cys Lys Lys His Tyr Trp Val Phe His Leu Phe Ser Trp Tyr
                470                 475                 480
Tyr Ile Phe Val Gln
                485


&lt;210&gt; 10
&lt;211&gt; 447
&lt;212&gt; PRT
&lt;213&gt; Homo sapiens

&lt;220&gt;
&lt;221&gt; misc_feature
&lt;223&gt; Incyte ID No:  1494955CD1

&lt;400&gt; 10
Met Glu Leu Ser Gln Met Ser Glu Leu Met Gly Leu Ser Val Leu
  1               5                  10                  15
Leu Gly Leu Leu Ala Leu Met Ala Thr Ala Ala Val Ala Arg Gly
                 20                  25                  30
Trp Leu Arg Ala Gly Glu Glu Arg Ser Gly Arg Pro Ala Cys Gln
                 35                  40                  45
Lys Ala Asn Gly Phe Pro Pro Asp Lys Ser Ser Gly Ser Lys Lys
                 50                  55                  60
Gln Lys Gln Tyr Gln Arg Ile Arg Lys Glu Lys Pro Gln Gln His
                 65                  70                  75
Asn Phe Thr His Arg Leu Leu Ala Ala Ala Leu Lys Ser His Ser
                 80                  85                  90
Gly Asn Ile Ser Cys Met Asp Phe Ser Ser Asn Gly Lys Tyr Leu
                 95                 100                 105
Ala Thr Cys Ala Asp Asp Arg Thr Ile Arg Ile Trp Ser Thr Lys
                110                 115                 120
Asp Phe Leu Gln Arg Glu His Arg Ser Met Arg Ala Asn Val Glu
                125                 130                 135
Leu Asp His Ala Thr Leu Val Arg Phe Ser Pro Asp Cys Arg Ala
                140                 145                 150
Phe Ile Val Trp Leu Ala Asn Gly Asp Thr Leu Arg Val Phe Lys
                155                 160                 165
Met Thr Lys Arg Glu Asp Gly Gly Tyr Thr Phe Thr Ala Thr Pro
                170                 175                 180
Glu Asp Phe Pro Lys Lys His Lys Ala Pro Val Ile Asp Ile Gly
```

```
                185                 190                 195
Ile Ala Asn Thr Gly Lys Phe Ile Met Thr Ala Ser Ser Asp Thr
                200                 205                 210
Thr Val Leu Ile Trp Ser Leu Lys Gly Gln Val Leu Ser Thr Ile
                215                 220                 225
Asn Thr Asn Gln Met Asn Asn Thr His Ala Ala Val Ser Pro Cys
                230                 235                 240
Gly Arg Phe Val Ala Ser Cys Gly Phe Thr Pro Asp Val Lys Val
                245                 250                 255
Trp Glu Val Cys Phe Gly Lys Lys Gly Glu Phe Gln Glu Val Val
                260                 265                 270
Arg Ala Phe Glu Leu Lys Gly His Ser Ala Ala Val His Ser Phe
                275                 280                 285
Ala Phe Ser Asn Asp Ser Arg Arg Met Ala Ser Val Ser Lys Asp
                290                 295                 300
Gly Thr Trp Lys Leu Trp Asp Thr Asp Val Glu Tyr Lys Lys Lys
                305                 310                 315
Gln Asp Pro Tyr Leu Leu Lys Thr Gly Arg Phe Glu Glu Ala Ala
                320                 325                 330
Gly Ala Ala Pro Cys Arg Leu Ala Leu Ser Pro Asn Ala Gln Val
                335                 340                 345
Leu Ala Leu Ala Ser Gly Ser Ser Ile His Leu Tyr Asn Thr Arg
                350                 355                 360
Arg Gly Glu Lys Glu Glu Cys Phe Glu Arg Val His Gly Glu Cys
                365                 370                 375
Ile Ala Asn Leu Ser Phe Asp Ile Thr Gly Arg Phe Leu Ala Ser
                380                 385                 390
Cys Gly Asp Arg Ala Val Arg Leu Phe His Asn Thr Pro Gly His
                395                 400                 405
Arg Ala Met Val Glu Glu Met Gln Gly His Leu Lys Arg Ala Ser
                410                 415                 420
Asn Glu Ser Thr Arg Gln Arg Leu Gln Gln Gln Leu Thr Gln Ala
                425                 430                 435
Gln Glu Thr Leu Lys Ser Leu Gly Ala Leu Lys Lys
                440                 445


<210> 11
<211> 199
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No:  1508161CD1

<400> 11
Met Pro Val Lys Lys Lys His Arg Ala Arg Met Ile Glu Tyr Phe
  1               5                  10                  15
Ile Asp Val Ala Arg Glu Cys Phe Asn Ile Gly Asn Phe Asn Ser
                 20                  25                  30
Leu Met Ala Ile Ile Ser Gly Met Asn Met Ser Pro Val Ser Arg
                 35                  40                  45
Leu Lys Lys Thr Trp Ala Lys Val Lys Thr Ala Lys Phe Asp Ile
                 50                  55                  60
```

```
Leu Glu His Gln Met Asp Pro Ser Ser Asn Phe Tyr Asn Tyr Arg
                65                  70                  75
Thr Ala Leu Arg Gly Ala Ala Gln Arg Ser Leu Thr Ala His Ser
                80                  85                  90
Ser Arg Glu Lys Ile Val Ile Pro Phe Phe Ser Leu Leu Ile Lys
                95                 100                 105
Asp Ile Tyr Phe Leu Asn Glu Gly Cys Ala Asn Arg Leu Pro Asn
               110                 115                 120
Gly His Val Asn Phe Glu Lys Phe Trp Glu Leu Ala Lys Gln Val
               125                 130                 135
Ser Glu Phe Met Thr Trp Lys Gln Val Glu Cys Pro Phe Glu Arg
               140                 145                 150
Asp Arg Lys Ile Leu Gln Tyr Leu Leu Thr Val Pro Val Phe Ser
               155                 160                 165
Glu Asp Ala Leu Tyr Leu Ala Ser Tyr Glu Ser Glu Gly Pro Glu
               170                 175                 180
Asn His Ile Glu Lys Asp Arg Trp Lys Ser Leu Arg Ser Ser Leu
               185                 190                 195
Leu Gly Arg Val


<210> 12
<211> 694
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No:  1811877CD1

<400> 12
Met Ala Phe Asp Pro Thr Ser Thr Leu Leu Ala Thr Gly Gly Cys
  1               5                  10                  15
Asp Gly Ala Val Arg Val Trp Asp Ile Val Arg His Tyr Gly Thr
                 20                  25                  30
His His Phe Arg Gly Ser Pro Gly Val Val His Leu Val Ala Phe
                 35                  40                  45
His Pro Asp Pro Thr Arg Leu Leu Leu Phe Ser Ser Ala Thr Asp
                 50                  55                  60
Ala Ala Ile Arg Val Trp Ser Leu Gln Asp Arg Ser Cys Leu Ala
                 65                  70                  75
Val Leu Thr Ala His Tyr Ser Ala Val Thr Ser Leu Ala Phe Ser
                 80                  85                  90
Ala Asp Gly His Thr Met Leu Ser Ser Gly Arg Asp Lys Ile Cys
                 95                 100                 105
Ile Ile Trp Asp Leu Gln Ser Cys Gln Ala Thr Arg Thr Val Pro
                110                 115                 120
Val Phe Glu Ser Val Glu Ala Ala Val Leu Leu Pro Glu Glu Pro
                125                 130                 135
Val Ser Gln Leu Gly Val Lys Ser Pro Gly Leu Tyr Phe Leu Thr
                140                 145                 150
Ala Gly Asp Gln Gly Thr Leu Arg Val Trp Glu Ala Ala Ser Gly
                155                 160                 165
Gln Cys Val Tyr Thr Gln Ala Gln Pro Pro Gly Pro Gly Gln Glu
                170                 175                 180
```

Leu Thr His Cys Thr Leu Ala His Thr Ala Gly Val Val Leu Thr
185 190 195
Ala Thr Ala Asp His Asn Leu Leu Leu Tyr Glu Ala Arg Ser Leu
200 205 210
Arg Leu Gln Lys Gln Phe Ala Gly Tyr Ser Glu Glu Val Leu Asp
215 220 225
Val Arg Phe Leu Gly Pro Glu Asp Ser His Val Val Val Ala Ser
230 235 240
Asn Ser Pro Cys Leu Lys Val Phe Glu Leu Gln Thr Ser Ala Cys
245 250 255
Gln Ile Leu His Gly His Thr Asp Ile Val Leu Ala Leu Asp Val
260 265 270
Phe Arg Lys Gly Trp Leu Phe Ala Ser Cys Ala Lys Asp Gln Ser
275 280 285
Val Arg Ile Trp Arg Met Asn Lys Ala Gly Gln Val Met Cys Val
290 295 300
Ala Gln Gly Ser Gly His Thr His Ser Val Gly Thr Val Cys Cys
305 310 315
Ser Arg Leu Lys Glu Ser Phe Leu Val Thr Gly Ser Gln Asp Cys
320 325 330
Thr Val Lys Leu Trp Pro Leu Pro Lys Ala Leu Leu Ser Lys Asn
335 340 345
Thr Ala Pro Asp Asn Gly Pro Ile Leu Leu Gln Ala Gln Thr Thr
350 355 360
Gln Arg Cys His Asp Lys Asp Ile Asn Ser Val Ala Ile Ala Pro
365 370 375
Asn Asp Lys Leu Leu Ala Thr Gly Ser Gln Asp Arg Thr Ala Lys
380 385 390
Leu Trp Ala Leu Pro Gln Cys Gln Leu Leu Gly Val Phe Ser Gly
395 400 405
His Arg Arg Gly Leu Trp Cys Val Gln Phe Ser Pro Met Asp Gln
410 415 420
Val Leu Ala Thr Ala Ser Ala Asp Gly Thr Ile Lys Leu Trp Ala
425 430 435
Leu Gln Asp Phe Ser Cys Leu Lys Thr Phe Glu Gly His Asp Ala
440 445 450
Ser Val Leu Lys Val Ala Phe Val Ser Arg Gly Thr Gln Leu Leu
455 460 465
Ser Ser Gly Ser Asp Gly Leu Val Lys Leu Trp Thr Ile Lys Asn
470 475 480
Asn Glu Cys Val Arg Thr Leu Asp Ala His Glu Asp Lys Val Trp
485 490 495
Gly Leu His Cys Ser Arg Leu Asp Asp His Ala Leu Thr Gly Ala
500 505 510
Ser Asp Ser Arg Val Ile Leu Trp Lys Asp Val Thr Glu Ala Glu
515 520 525
Gln Ala Glu Glu Gln Ala Arg Gln Glu Glu Gln Val Val Arg Gln
530 535 540
Gln Glu Leu Asp Asn Leu Leu His Glu Lys Arg Tyr Leu Arg Ala
545 550 555
Leu Gly Leu Ala Ile Ser Leu Asp Arg Pro His Thr Val Leu Thr
560 565 570
Val Ile Gln Ala Ile Arg Arg Asp Pro Glu Ala Cys Glu Lys Leu
575 580 585
Glu Ala Thr Met Leu Arg Leu Arg Arg Asp Gln Lys Glu Ala Leu

```
                590                 595                 600
Leu Arg Phe Cys Val Thr Trp Asn Thr Asn Ser Arg His Cys His
                605                 610                 615
Glu Ala Gln Ala Val Leu Gly Val Leu Leu Arg Arg Glu Ala Pro
                620                 625                 630
Glu Glu Leu Leu Ala Tyr Glu Gly Val Arg Ala Ala Leu Glu Ala
                635                 640                 645
Leu Leu Pro Tyr Thr Glu Arg His Phe Gln Arg Leu Ser Arg Thr
                650                 655                 660
Leu Gln Ala Ala Ala Phe Leu Asp Phe Leu Trp His Asn Met Lys
                665                 670                 675
Leu Pro Val Pro Ala Ala Ala Pro Thr Pro Trp Glu Thr His Lys
                680                 685                 690
Gly Ala Leu Pro


<210> 13
<211> 654
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No:  1848674CD1

<400> 13
Met Glu Arg Ser Gly Pro Ser Glu Val Thr Gly Ser Asp Ala Ser
  1               5                  10                  15
Gly Pro Asp Pro Gln Leu Ala Val Thr Met Gly Phe Thr Gly Phe
                 20                  25                  30
Gly Lys Lys Ala Arg Thr Phe Asp Leu Glu Ala Met Phe Glu Gln
                 35                  40                  45
Thr Arg Arg Thr Ala Val Glu Arg Ser Arg Lys Thr Leu Glu Ala
                 50                  55                  60
Arg Glu Lys Glu Glu Glu Met Asn Arg Glu Lys Glu Leu Arg Arg
                 65                  70                  75
Gln Asn Glu Asp Ile Glu Pro Thr Ser Ser Arg Ser Asn Val Val
                 80                  85                  90
Arg Asp Cys Ser Lys Ser Ser Ser Arg Asp Thr Ser Ser Ser Glu
                 95                 100                 105
Ser Glu Gln Ser Ser Asp Ser Ser Asp Asp Glu Leu Ile Gly Pro
                110                 115                 120
Pro Leu Pro Pro Lys Met Val Gly Lys Pro Val Asn Phe Met Glu
                125                 130                 135
Glu Asp Ile Leu Gly Pro Leu Pro Pro Pro Leu Asn Glu Glu Glu
                140                 145                 150
Glu Glu Ala Glu Glu Glu Glu Glu Glu Glu Glu Glu Glu Glu Asn
                155                 160                 165
Pro Val His Lys Ile Pro Asp Ser His Glu Ile Thr Leu Lys His
                170                 175                 180
Gly Thr Lys Thr Val Ser Ala Leu Gly Leu Asp Pro Ser Gly Ala
                185                 190                 195
Arg Leu Val Thr Gly Gly Tyr Asp Tyr Asp Val Lys Phe Trp Asp
                200                 205                 210
Phe Ala Gly Met Asp Ala Ser Phe Lys Ala Phe Arg Ser Leu Gln
```

```
                215                 220                 225
Pro Cys Glu Cys His Gln Ile Lys Ser Leu Gln Tyr Ser Asn Thr
                230                 235                 240
Gly Asp Met Ile Leu Val Val Ser Gly Ser Ser Gln Ala Lys Val
                245                 250                 255
Ile Asp Arg Asp Gly Phe Glu Val Met Glu Cys Ile Lys Gly Asp
                260                 265                 270
Gln Tyr Ile Val Asp Met Ala Asn Thr Lys Gly His Thr Ala Met
                275                 280                 285
Leu His Thr Gly Ser Trp His Pro Lys Ile Lys Gly Glu Phe Met
                290                 295                 300
Thr Cys Ser Asn Asp Ala Thr Val Arg Thr Trp Glu Val Glu Asn
                305                 310                 315
Pro Lys Lys Gln Lys Ser Val Phe Lys Pro Arg Thr Met Gln Gly
                320                 325                 330
Lys Lys Val Ile Pro Thr Thr Cys Thr Tyr Ser Arg Asp Gly Asn
                335                 340                 345
Leu Ile Ala Ala Ala Cys Gln Asn Gly Ser Ile Gln Ile Trp Asp
                350                 355                 360
Arg Asn Leu Thr Val His Pro Lys Phe His Tyr Lys Gln Ala His
                365                 370                 375
Asp Ser Gly Thr Asp Thr Ser Cys Val Thr Phe Ser Tyr Asp Gly
                380                 385                 390
Asn Val Leu Ala Ser Arg Gly Gly Asp Asp Ser Leu Lys Leu Trp
                395                 400                 405
Asp Ile Arg Gln Phe Asn Lys Pro Leu Phe Ser Ala Ser Gly Leu
                410                 415                 420
Pro Thr Met Phe Pro Met Thr Asp Cys Cys Phe Ser Pro Asp Asp
                425                 430                 435
Lys Leu Ile Val Thr Gly Thr Ser Ile Gln Arg Gly Cys Gly Ser
                440                 445                 450
Gly Lys Leu Val Phe Phe Glu Arg Arg Thr Phe Gln Arg Val Tyr
                455                 460                 465
Glu Ile Asp Ile Thr Asp Ala Ser Val Val Arg Cys Leu Trp His
                470                 475                 480
Pro Lys Leu Asn Gln Ile Met Val Gly Thr Gly Asn Gly Leu Ala
                485                 490                 495
Lys Val Tyr Tyr Asp Pro Asn Lys Ser Gln Arg Gly Ala Lys Leu
                500                 505                 510
Cys Val Val Lys Thr Gln Arg Lys Ala Lys Gln Ala Glu Thr Leu
                515                 520                 525
Thr Gln Asp Tyr Ile Ile Thr Pro His Ala Leu Pro Met Phe Arg
                530                 535                 540
Glu Pro Arg Gln Arg Ser Thr Arg Lys Gln Leu Glu Lys Asp Arg
                545                 550                 555
Leu Asp Pro Leu Lys Ser His Lys Pro Glu Pro Pro Val Ala Gly
                560                 565                 570
Pro Gly Arg Gly Gly Arg Val Gly Thr His Gly Gly Thr Leu Ser
                575                 580                 585
Ser Tyr Ile Val Lys Asn Ile Ala Leu Asp Lys Thr Asp Asp Ser
                590                 595                 600
Asn Pro Arg Glu Ala Ile Leu Arg His Ala Lys Ala Ala Glu Asp
                605                 610                 615
Ser Pro Tyr Trp Val Ser Pro Ala Tyr Ser Lys Thr Gln Pro Lys
                620                 625                 630
```

```
Thr Met Phe Ala Gln Val Glu Ser Asp Asp Glu Glu Ala Lys Asn
                635                 640                 645
Glu Pro Glu Trp Lys Lys Arg Lys Ile
                650


<210> 14
<211> 180
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No:  2012970CD1

<400> 14
Met Glu Ala Asn Met Pro Lys Arg Lys Glu Pro Gly Arg Ser Leu
  1               5                  10                  15
Arg Ile Lys Val Ile Ser Met Gly Asn Ala Glu Val Gly Lys Ser
                 20                  25                  30
Cys Ile Ile Lys Arg Tyr Cys Glu Lys Arg Phe Val Ser Lys Tyr
                 35                  40                  45
Leu Ala Thr Ile Gly Ile Asp Tyr Gly Val Thr Lys Val His Val
                 50                  55                  60
Arg Asp Arg Glu Ile Lys Val Asn Ile Phe Asp Met Ala Gly His
                 65                  70                  75
Pro Phe Phe Tyr Glu Val Arg Asn Glu Phe Tyr Lys Asp Thr Gln
                 80                  85                  90
Gly Val Ile Leu Val Tyr Asp Val Gly Gln Lys Asp Ser Phe Asp
                 95                 100                 105
Ala Leu Asp Ala Trp Leu Ala Glu Met Lys Gln Glu Leu Gly Pro
                110                 115                 120
His Gly Asn Met Glu Asn Ile Ile Phe Val Val Cys Ala Asn Lys
                125                 130                 135
Ile Asp Cys Thr Lys His Arg Cys Val Asp Glu Ser Glu Gly Arg
                140                 145                 150
Leu Trp Ala Glu Ser Lys Gly Phe Leu Tyr Phe Glu Thr Ser Ala
                155                 160                 165
Gln Thr Gly Glu Gly Ile Asn Glu Met Phe Gln Ile His Leu Gly
                170                 175                 180


<210> 15
<211> 374
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No:  2254315CD1

<400> 15
Met Ala Ala Ser Ala Ala Ala Ala Glu Leu Gln Ala Ser Gly Gly
  1               5                  10                  15
Pro Arg His Pro Val Cys Leu Leu Val Leu Gly Met Ala Gly Ser
```

```
                20                  25                  30
Gly Lys Thr Thr Phe Val Gln Arg Leu Thr Gly His Leu His Ala
                35                  40                  45
Gln Gly Thr Pro Pro Tyr Val Ile Asn Leu Asp Pro Ala Val His
                50                  55                  60
Glu Val Pro Phe Pro Ala Asn Ile Asp Ile Arg Asp Thr Val Lys
                65                  70                  75
Tyr Lys Glu Val Met Lys Gln Tyr Gly Leu Gly Pro Asn Gly Gly
                80                  85                  90
Ile Val Thr Ser Leu Asn Leu Phe Ala Thr Arg Phe Asp Gln Val
                95                 100                 105
Met Lys Phe Ile Glu Lys Ala Gln Asn Met Ser Lys Tyr Val Leu
               110                 115                 120
Ile Asp Thr Pro Gly Gln Ile Glu Val Phe Thr Trp Ser Ala Ser
               125                 130                 135
Gly Thr Ile Ile Thr Glu Ala Leu Ala Ser Ser Phe Pro Thr Val
               140                 145                 150
Val Ile Tyr Val Met Asp Thr Ser Arg Ser Thr Asn Pro Val Thr
               155                 160                 165
Phe Met Ser Asn Met Leu Tyr Ala Cys Ser Ile Leu Tyr Lys Thr
               170                 175                 180
Lys Leu Pro Phe Ile Val Val Met Asn Lys Thr Asp Ile Ile Asp
               185                 190                 195
His Ser Phe Ala Val Glu Trp Met Gln Asp Phe Glu Ala Phe Gln
               200                 205                 210
Asp Ala Leu Asn Gln Glu Thr Thr Tyr Val Ser Asn Leu Thr Arg
               215                 220                 225
Ser Met Ser Leu Val Leu Asp Glu Phe Tyr Ser Ser Leu Arg Val
               230                 235                 240
Val Gly Val Ser Ala Val Leu Gly Thr Gly Leu Asp Glu Leu Phe
               245                 250                 255
Val Gln Val Thr Ser Ala Ala Glu Glu Tyr Glu Arg Glu Tyr Arg
               260                 265                 270
Pro Glu Tyr Glu Arg Leu Lys Lys Ser Leu Ala Asn Ala Glu Ser
               275                 280                 285
Gln Gln Gln Arg Glu Gln Leu Glu Arg Leu Arg Lys Asp Met Gly
               290                 295                 300
Ser Val Ala Leu Asp Ala Gly Thr Ala Lys Asp Ser Leu Ser Pro
               305                 310                 315
Val Leu His Pro Ser Asp Leu Ile Leu Thr Arg Gly Thr Leu Asp
               320                 325                 330
Glu Glu Asp Glu Glu Ala Asp Ser Asp Thr Asp Asp Ile Asp His
               335                 340                 345
Arg Val Thr Glu Glu Ser His Glu Glu Pro Ala Phe Gln Asn Phe
               350                 355                 360
Met Gln Glu Ser Met Ala Gln Tyr Trp Lys Arg Asn Asn Lys
               365                 370
```

<210> 16
<211> 649
<212> PRT
<213> Homo sapiens

<220>

<221> misc_feature
<223> Incyte ID No: 2415545CD1

<400> 16
Met Glu Gly Ala Gly Tyr Arg Val Val Phe Glu Lys Gly Gly Val
1 5 10 15
Tyr Leu His Thr Ser Ala Lys Lys Tyr Gln Asp Arg Asp Ser Leu
20 25 30
Ile Ala Gly Val Ile Arg Val Val Glu Lys Asp Asn Asp Val Leu
35 40 45
Leu His Trp Ala Pro Val Glu Glu Ala Gly Asp Ser Thr Gln Ile
50 55 60
Leu Phe Ser Lys Lys Asp Ser Ser Gly Gly Asp Ser Cys Ala Ser
65 70 75
Glu Glu Glu Pro Thr Phe Asp Pro Gly Tyr Glu Pro Asp Trp Ala
80 85 90
Val Ile Ser Thr Val Arg Pro Gln Pro Cys His Ser Glu Pro Thr
95 100 105
Arg Gly Ala Glu Pro Ser Cys Pro Gln Gly Ser Trp Ala Phe Ser
110 115 120
Val Ser Leu Gly Glu Leu Lys Ser Ile Arg Arg Ser Lys Pro Gly
125 130 135
Leu Ser Trp Ala Tyr Leu Val Leu Val Thr Gln Ala Gly Gly Ser
140 145 150
Leu Pro Ala Leu His Phe His Arg Gly Gly Thr Arg Ala Leu Leu
155 160 165
Arg Val Leu Ser Arg Tyr Leu Leu Leu Ala Ser Ser Pro Gln Asp
170 175 180
Ser Arg Leu Tyr Leu Val Phe Pro His Asp Ser Ser Ala Leu Ser
185 190 195
Asn Ser Phe His His Leu Gln Leu Phe Asp Gln Asp Ser Ser Asn
200 205 210
Val Val Ser Arg Phe Leu Gln Asp Pro Tyr Ser Thr Thr Phe Ser
215 220 225
Ser Phe Ser Arg Val Thr Asn Phe Phe Arg Gly Ala Leu Gln Pro
230 235 240
Gln Pro Glu Gly Ala Ala Ser Asp Leu Pro Pro Pro Pro Asp Asp
245 250 255
Glu Pro Glu Pro Gly Phe Glu Val Ile Ser Cys Val Glu Leu Gly
260 265 270
Pro Arg Pro Thr Val Glu Arg Gly Pro Pro Val Thr Glu Glu Glu
275 280 285
Trp Ala Arg His Val Gly Pro Glu Gly Arg Leu Gln Gln Val Pro
290 295 300
Glu Leu Lys Asn Arg Ile Phe Ser Gly Gly Leu Ser Pro Ser Leu
305 310 315
Arg Arg Glu Ala Trp Lys Phe Leu Leu Gly Tyr Leu Ser Trp Glu
320 325 330
Gly Thr Ala Glu Glu His Lys Ala His Ile Arg Lys Lys Thr Asp
335 340 345
Glu Tyr Phe Arg Met Lys Leu Gln Trp Lys Ser Val Ser Pro Glu
350 355 360
Gln Glu Arg Arg Asn Ser Leu Leu His Gly Tyr Arg Ser Leu Ile
365 370 375
Glu Arg Asp Val Ser Arg Thr Asp Arg Thr Asn Lys Phe Tyr Glu

380 385 390
Gly Pro Glu Asn Pro Gly Leu Gly Leu Leu Asn Asp Ile Leu Leu
395 400 405
Thr Tyr Cys Met Tyr His Phe Asp Leu Gly Tyr Val Gln Gly Met
410 415 420
Ser Asp Leu Leu Ser Pro Ile Leu Tyr Val Ile Gln Asn Glu Val
425 430 435
Asp Ala Phe Trp Cys Phe Cys Gly Phe Met Glu Leu Val Gln Gly
440 445 450
Asn Phe Glu Glu Ser Gln Glu Thr Met Lys Arg Gln Leu Gly Arg
455 460 465
Leu Leu Leu Leu Leu Arg Val Leu Asp Pro Leu Leu Cys Asp Phe
470 475 480
Leu Asp Ser Gln Asp Ser Gly Ser Leu Cys Phe Cys Phe Arg Trp
485 490 495
Leu Leu Ile Trp Phe Lys Arg Glu Phe Pro Phe Pro Asp Val Leu
500 505 510
Arg Leu Trp Glu Val Leu Trp Thr Gly Leu Pro Gly Pro Asn Leu
515 520 525
His Leu Leu Val Ala Cys Ala Ile Leu Asp Met Glu Arg Asp Thr
530 535 540
Leu Met Leu Ser Gly Phe Gly Ser Asn Glu Ile Leu Lys His Ile
545 550 555
Asn Glu Leu Thr Met Lys Leu Ser Val Glu Asp Val Leu Thr Arg
560 565 570
Ala Glu Ala Leu His Arg Gln Leu Thr Ala Cys Thr Arg Ala Ala
575 580 585
Pro Gln Arg Ala Gly Asp Pro Gly Ala Gly Pro Ala Thr Gln Ser
590 595 600
Pro Thr Ala Pro Arg Pro Pro Pro Pro Arg Cys Leu Cys Thr Pro
605 610 615
Thr Arg Ala Pro Pro Thr Pro Pro Pro Ser Thr Asp Thr Ala Pro
620 625 630
Gln Pro Asp Ser Ser Leu Glu Ile Leu Pro Glu Glu Glu Asp Glu
635 640 645
Gly Ala Asp Ser

<210> 17
<211> 698
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2707969CD1

<400> 17
Met Cys His Asp Asp Asp Asp Lys Asp Pro Val Leu Arg Val Phe
1 5 10 15
Asp Ser Arg Val Asp Lys Ile Arg Leu Leu Asn Val Arg Thr Pro
20 25 30
Thr Leu Arg Thr Ser Met Tyr Gln Lys Cys Thr Thr Val Asp Glu
35 40 45
Ala Glu Lys Ala Ile Glu Leu Arg Leu Ala Lys Ile Asp His Thr

```
                50                  55                  60
Ala Ile His Pro His Leu Leu Asp Met Lys Ile Gly Gln Gly Lys
                65                  70                  75
Tyr Glu Pro Gly Phe Phe Pro Lys Leu Gln Ser Asp Val Leu Ser
                80                  85                  90
Thr Gly Pro Ala Ser Asn Lys Trp Thr Lys Arg Asn Ala Pro Ala
                95                  100                 105
Gln Trp Arg Arg Lys Asp Arg Gln Lys Gln His Thr Glu His Leu
                110                 115                 120
Arg Leu Asp Asn Asp Gln Arg Glu Lys Tyr Ile Gln Glu Ala Arg
                125                 130                 135
Thr Met Gly Ser Thr Ile Arg Gln Pro Lys Leu Ser Asn Leu Ser
                140                 145                 150
Pro Ser Val Ile Ala Gln Thr Asn Trp Lys Phe Val Glu Gly Leu
                155                 160                 165
Leu Lys Glu Cys Arg Asn Lys Thr Lys Arg Met Leu Val Glu Lys
                170                 175                 180
Met Gly Arg Glu Ala Val Glu Leu Gly His Gly Glu Val Asn Ile
                185                 190                 195
Thr Gly Val Glu Glu Asn Thr Leu Ile Ala Ser Leu Cys Asp Leu
                200                 205                 210
Leu Glu Arg Ile Trp Ser His Gly Leu Gln Val Lys Gln Gly Lys
                215                 220                 225
Ser Ala Leu Trp Ser His Leu Leu His Tyr Gln Asp Asn Arg Gln
                230                 235                 240
Arg Lys Leu Thr Ser Gly Ser Leu Ser Thr Ser Gly Ile Leu Leu
                245                 250                 255
Asp Ser Glu Arg Arg Lys Ser Asp Ala Ser Ser Leu Met Pro Pro
                260                 265                 270
Leu Arg Ile Ser Leu Ile Gln Asp Met Arg His Ile Gln Asn Ile
                275                 280                 285
Gly Glu Ile Lys Thr Asp Val Gly Lys Ala Arg Ala Trp Val Arg
                290                 295                 300
Leu Ser Met Glu Lys Lys Leu Leu Ser Arg His Leu Lys Gln Leu
                305                 310                 315
Leu Ser Asp His Glu Leu Thr Lys Lys Leu Tyr Lys Arg Tyr Ala
                320                 325                 330
Phe Leu Arg Cys Asp Asp Glu Lys Glu Gln Phe Leu Tyr His Leu
                335                 340                 345
Leu Ser Phe Asn Ala Val Asp Tyr Phe Cys Phe Thr Asn Val Phe
                350                 355                 360
Thr Thr Ile Leu Ile Pro Tyr His Ile Leu Ile Val Pro Ser Lys
                365                 370                 375
Lys Leu Gly Gly Ser Met Phe Thr Ala Asn Pro Trp Ile Cys Ile
                380                 385                 390
Ser Gly Glu Leu Gly Glu Thr Gln Ile Met Gln Ile Pro Arg Asn
                395                 400                 405
Val Leu Glu Met Thr Phe Glu Cys Gln Asn Leu Gly Lys Leu Thr
                410                 415                 420
Thr Val Gln Ile Gly His Asp Asn Ser Gly Leu Tyr Ala Lys Trp
                425                 430                 435
Leu Val Glu Tyr Val Met Val Arg Asn Glu Ile Thr Gly His Thr
                440                 445                 450
Tyr Lys Phe Pro Cys Gly Arg Trp Leu Gly Lys Gly Met Asp Asp
                455                 460                 465
```

```
Gly Ser Leu Glu Arg Ile Leu Val Gly Glu Leu Leu Thr Ser Gln
                470                 475                 480
Pro Glu Val Asp Glu Arg Pro Cys Arg Thr Pro Pro Leu Gln Gln
                485                 490                 495
Ser Pro Ser Val Ile Arg Arg Leu Val Thr Ile Ser Pro Asn Asn
                500                 505                 510
Lys Pro Lys Leu Asn Thr Gly Gln Ile Gln Glu Ser Ile Gly Glu
                515                 520                 525
Ala Val Asn Gly Ile Val Lys His Phe His Lys Pro Glu Lys Glu
                530                 535                 540
Arg Gly Ser Leu Thr Leu Leu Leu Cys Gly Glu Cys Gly Leu Val
                545                 550                 555
Ser Ala Leu Glu Gln Ala Phe Gln His Gly Phe Lys Ser Pro Arg
                560                 565                 570
Leu Phe Lys Asn Val Phe Ile Trp Asp Phe Leu Glu Lys Ala Gln
                575                 580                 585
Thr Tyr Tyr Glu Thr Leu Glu Lys Asn Glu Val Val Pro Glu Glu
                590                 595                 600
Asn Trp His Thr Arg Ala Arg Asn Phe Cys Arg Phe Val Thr Ala
                605                 610                 615
Ile Asn Asn Thr Pro Arg Asn Ile Gly Lys Asp Gly Lys Phe Gln
                620                 625                 630
Met Leu Val Cys Leu Gly Ala Arg Asp His Leu Leu His His Trp
                635                 640                 645
Ile Ala Leu Leu Ala Asp Cys Pro Ile Thr Ala His Met Tyr Glu
                650                 655                 660
Asp Val Ala Leu Ile Lys Asp His Thr Leu Val Asn Ser Leu Ile
                665                 670                 675
Arg Val Leu Gln Thr Leu Gln Glu Phe Asn Ile Thr Leu Glu Thr
                680                 685                 690
Ser Leu Val Lys Gly Ile Asp Ile
                695

<210> 18
<211> 396
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No:  2817769CD1

<400> 18
Met Pro Pro Lys Lys Gly Gly Asp Gly Ile Lys Pro Pro Pro Ile
  1               5                  10                  15
Ile Gly Arg Phe Gly Thr Ser Leu Lys Ile Gly Ile Val Gly Leu
                 20                  25                  30
Pro Asn Val Gly Lys Ser Thr Phe Phe Asn Val Leu Thr Asn Ser
                 35                  40                  45
Gln Ala Ser Ala Glu Asn Phe Pro Phe Cys Thr Ile Asp Pro Asn
                 50                  55                  60
Glu Ser Arg Val Pro Val Pro Asp Glu Arg Phe Asp Phe Leu Cys
                 65                  70                  75
Gln Tyr His Lys Pro Ala Ser Lys Ile Pro Ala Phe Leu Asn Val
                 80                  85                  90
```

```
Val Asp Ile Ala Gly Leu Val Lys Gly Ala His Asn Gly Gln Gly
                95                  100                 105
Leu Gly Asn Ala Phe Leu Ser His Ile Ser Ala Cys Asp Gly Ile
                110                 115                 120
Phe His Leu Thr Arg Ala Phe Glu Asp Asp Asp Ile Thr His Val
                125                 130                 135
Glu Gly Ser Val Asp Pro Ile Arg Asp Ile Glu Ile Ile His Glu
                140                 145                 150
Glu Leu Gln Leu Lys Asp Glu Glu Met Ile Gly Pro Ile Ile Asp
                155                 160                 165
Lys Leu Glu Lys Val Ala Val Arg Gly Gly Asp Lys Lys Leu Lys
                170                 175                 180
Pro Glu Tyr Asp Ile Met Cys Lys Val Lys Ser Trp Val Ile Asp
                185                 190                 195
Gln Lys Lys Pro Val Arg Phe Tyr His Asp Trp Asn Asp Lys Glu
                200                 205                 210
Ile Glu Val Leu Asn Lys His Leu Phe Leu Thr Ser Lys Pro Met
                215                 220                 225
Val Tyr Leu Val Asn Leu Ser Glu Lys Asp Tyr Ile Arg Lys Lys
                230                 235                 240
Asn Lys Trp Leu Ile Lys Ile Lys Glu Trp Val Asp Lys Tyr Asp
                245                 250                 255
Pro Gly Ala Leu Val Ile Pro Phe Ser Gly Ala Leu Glu Leu Lys
                260                 265                 270
Leu Gln Glu Leu Ser Ala Glu Glu Arg Gln Lys Tyr Leu Glu Ala
                275                 280                 285
Asn Met Thr Gln Ser Ala Leu Pro Lys Ile Ile Lys Ala Gly Phe
                290                 295                 300
Ala Ala Leu Gln Leu Glu Tyr Phe Phe Thr Ala Gly Pro Asp Glu
                305                 310                 315
Val Arg Ala Trp Thr Ile Arg Lys Gly Thr Lys Ala Pro Gln Ala
                320                 325                 330
Ala Gly Lys Ile His Thr Asp Phe Glu Lys Gly Phe Ile Met Ala
                335                 340                 345
Glu Val Met Lys Tyr Glu Asp Phe Lys Glu Glu Gly Ser Glu Asn
                350                 355                 360
Ala Val Lys Ala Ala Gly Lys Tyr Arg Gln Gln Gly Arg Asn Tyr
                365                 370                 375
Ile Val Glu Asp Gly Asp Ile Ile Phe Phe Lys Phe Asn Thr Pro
                380                 385                 390
Gln Gln Pro Lys Lys Lys
                395


<210> 19
<211> 634
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No:  2917557CD1

<400> 19
Met Ser Ser Asp Ser Glu Tyr Asp Ser Asp Asp Asp Arg Thr Lys
```

```
1               5                   10                  15
Glu Glu Arg Ala Tyr Asp Lys Ala Lys Arg Arg Ile Glu Lys Arg
                20                  25                  30
Arg Leu Glu His Ser Lys Asn Val Asn Thr Glu Lys Leu Arg Ala
                35                  40                  45
Pro Ile Ile Cys Val Leu Gly His Val Asp Thr Gly Lys Thr Lys
                50                  55                  60
Ile Leu Asp Lys Leu Arg His Thr His Val Gln Asp Gly Glu Ala
                65                  70                  75
Gly Gly Ile Thr Gln Gln Ile Gly Ala Thr Asn Val Pro Leu Glu
                80                  85                  90
Ala Ile Asn Glu Gln Thr Lys Met Ile Lys Asn Phe Asp Arg Glu
                95                  100                 105
Asn Val Arg Ile Pro Gly Met Leu Ile Ile Asp Thr Pro Gly His
                110                 115                 120
Glu Ser Phe Ser Asn Leu Arg Asn Arg Gly Ser Ser Leu Cys Asp
                125                 130                 135
Ile Ala Ile Leu Val Val Asp Ile Met His Gly Leu Glu Pro Gln
                140                 145                 150
Thr Ile Glu Ser Ile Asn Leu Leu Lys Ser Lys Lys Cys Pro Phe
                155                 160                 165
Ile Val Ala Leu Asn Lys Ile Asp Arg Leu Tyr Asp Trp Lys Lys
                170                 175                 180
Ser Pro Asp Ser Asp Val Ala Ala Thr Leu Lys Lys Gln Lys Lys
                185                 190                 195
Asn Thr Lys Asp Glu Phe Glu Glu Arg Ala Lys Ala Ile Ile Val
                200                 205                 210
Glu Phe Ala Gln Gln Gly Leu Asn Ala Ala Leu Phe Tyr Glu Asn
                215                 220                 225
Lys Asp Pro Arg Thr Phe Val Ser Leu Val Pro Thr Ser Ala His
                230                 235                 240
Thr Gly Asp Gly Met Gly Ser Leu Ile Tyr Leu Leu Val Glu Leu
                245                 250                 255
Thr Gln Thr Met Leu Ser Lys Arg Leu Ala His Cys Glu Glu Leu
                260                 265                 270
Arg Ala Gln Val Met Glu Val Lys Ala Leu Pro Gly Met Gly Thr
                275                 280                 285
Thr Ile Asp Val Ile Leu Ile Asn Gly Arg Leu Lys Glu Gly Asp
                290                 295                 300
Thr Ile Ile Val Pro Gly Val Glu Gly Pro Ile Val Thr Gln Ile
                305                 310                 315
Arg Gly Leu Leu Leu Pro Pro Pro Met Lys Glu Leu Arg Val Lys
                320                 325                 330
Asn Gln Tyr Glu Lys His Lys Glu Val Glu Ala Ala Gln Gly Val
                335                 340                 345
Lys Ile Leu Gly Lys Asp Leu Glu Lys Thr Leu Ala Gly Leu Pro
                350                 355                 360
Leu Leu Val Ala Tyr Lys Glu Asp Glu Ile Pro Val Leu Lys Asp
                365                 370                 375
Glu Leu Ile His Glu Leu Lys Gln Thr Leu Asn Ala Ile Lys Leu
                380                 385                 390
Glu Glu Lys Gly Val Tyr Val Gln Ala Ser Thr Leu Gly Ser Leu
                395                 400                 405
Glu Ala Leu Leu Glu Phe Leu Lys Thr Ser Glu Val Pro Tyr Ala
                410                 415                 420
```

Gly Ile Asn Ile Gly Pro Val His Lys Lys Asp Val Met Lys Ala
425 430 435
Ser Val Met Leu Glu His Asp Pro Gln Tyr Ala Val Ile Leu Ala
440 445 450
Phe Asp Val Arg Ile Glu Arg Asp Ala Gln Glu Met Ala Asp Ser
455 460 465
Leu Gly Val Arg Ile Phe Ser Ala Glu Ile Ile Tyr His Leu Phe
470 475 480
Asp Ala Phe Thr Lys Tyr Arg Gln Asp Tyr Lys Lys Gln Lys Gln
485 490 495
Glu Glu Phe Lys His Ile Ala Val Phe Pro Cys Lys Ile Lys Ile
500 505 510
Leu Pro Gln Tyr Ile Phe Asn Ser Arg Asp Pro Ile Val Met Gly
515 520 525
Val Thr Val Glu Ala Gly Gln Val Lys Gln Gly Thr Pro Met Cys
530 535 540
Val Pro Ser Lys Asn Phe Val Asp Ile Gly Ile Val Thr Ser Ile
545 550 555
Glu Ile Asn His Lys Gln Val Asp Val Ala Lys Lys Gly Gln Glu
560 565 570
Val Cys Val Lys Ile Glu Pro Ile Pro Gly Glu Ser Pro Lys Met
575 580 585
Phe Gly Arg His Phe Glu Ala Thr Asp Ile Leu Val Ser Lys Ile
590 595 600
Ser Arg Gln Ser Ile Asp Ala Leu Lys Asp Trp Phe Arg Asp Glu
605 610 615
Met Gln Lys Ser Asp Trp Gln Leu Ile Val Glu Leu Lys Lys Val
620 625 630
Phe Glu Ile Ile

<210> 20
<211> 196
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3421335CD1

<400> 20
Met Gly Ser Val Asn Ser Arg Gly His Lys Ala Glu Ala Gln Val
1 5 10 15
Val Met Met Gly Leu Asp Ser Ala Gly Lys Thr Thr Leu Leu Tyr
20 25 30
Lys Leu Lys Gly His Gln Leu Val Glu Thr Leu Pro Thr Val Gly
35 40 45
Phe Asn Val Glu Pro Leu Lys Ala Pro Gly His Val Ser Leu Thr
50 55 60
Leu Trp Asp Val Gly Gly Gln Ala Pro Leu Arg Ala Ser Trp Lys
65 70 75
Asp Tyr Leu Glu Gly Thr Asp Ile Leu Val Tyr Val Leu Asp Ser
80 85 90
Thr Asp Glu Ala Arg Leu Pro Glu Ser Ala Ala Glu Leu Thr Glu
95 100 105

```
Val Leu Asn Asp Pro Asn Met Ala Gly Val Pro Phe Leu Val Leu
                110                 115                 120
Ala Asn Lys Gln Glu Ala Pro Asp Ala Leu Pro Leu Leu Lys Ile
                125                 130                 135
Arg Asn Arg Leu Ser Leu Glu Arg Phe Gln Asp His Cys Trp Glu
                140                 145                 150
Leu Arg Gly Cys Ser Ala Leu Thr Gly Glu Gly Leu Pro Glu Ala
                155                 160                 165
Leu Gln Ser Leu Trp Ser Leu Leu Lys Ser Arg Ser Cys Met Cys
                170                 175                 180
Leu Gln Ala Arg Ala His Gly Ala Glu Arg Gly Asp Ser Lys Arg
                185                 190                 195
Ser


<210> 21
<211> 446
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No:  605761CD1

<400> 21
Met Ala Ala Arg Lys Gly Arg Arg Arg Thr Cys Glu Thr Gly Glu
  1             5                   10                  15
Pro Met Glu Ala Glu Ser Gly Asp Thr Ser Ser Glu Gly Pro Ala
                20                  25                  30
Gln Val Tyr Leu Pro Gly Arg Gly Pro Pro Leu Arg Glu Gly Glu
                35                  40                  45
Glu Leu Val Met Asp Glu Glu Ala Tyr Val Leu Tyr His Arg Ala
                50                  55                  60
Gln Thr Gly Ala Pro Cys Leu Ser Phe Asp Ile Val Arg Asp His
                65                  70                  75
Leu Gly Asp Asn Arg Thr Glu Leu Pro Leu Thr Leu Tyr Leu Cys
                80                  85                  90
Ala Gly Thr Gln Ala Glu Ser Ala Gln Ser Asn Arg Leu Met Met
                95                  100                 105
Leu Arg Met His Asn Leu His Gly Thr Lys Pro Pro Pro Ser Glu
                110                 115                 120
Gly Ser Asp Glu Glu Glu Glu Glu Glu Asp Glu Glu Asp Glu Glu
                125                 130                 135
Glu Arg Lys Pro Gln Leu Glu Leu Ala Met Val Pro His Tyr Gly
                140                 145                 150
Gly Ile Asn Arg Val Arg Val Ser Trp Leu Gly Glu Glu Pro Val
                155                 160                 165
Ala Gly Val Trp Ser Glu Lys Gly Gln Val Glu Val Phe Ala Leu
                170                 175                 180
Arg Arg Leu Leu Gln Val Val Glu Glu Pro Gln Ala Leu Ala Ala
                185                 190                 195
Phe Leu Arg Asp Glu Gln Ala Gln Met Lys Pro Ile Phe Ser Phe
                200                 205                 210
Ala Gly His Met Gly Glu Gly Phe Ala Leu Asp Trp Ser Pro Arg
                215                 220                 225
```

Val Thr Gly Arg Leu Leu Thr Gly Asp Cys Gln Lys Asn Ile His
230 235 240
Leu Trp Thr Pro Thr Asp Gly Gly Ser Trp His Val Asp Gln Arg
245 250 255
Pro Phe Val Gly His Thr Arg Ser Val Glu Asp Leu Gln Trp Ser
260 265 270
Pro Thr Glu Asn Thr Val Phe Ala Ser Cys Ser Ala Asp Ala Ser
275 280 285
Ile Arg Ile Trp Asp Ile Arg Ala Ala Pro Ser Lys Ala Cys Met
290 295 300
Leu Thr Thr Ala Thr Ala His Asp Gly Asp Val Asn Val Ile Ser
305 310 315
Trp Ser Arg Arg Glu Pro Phe Leu Leu Ser Gly Gly Asp Asp Gly
320 325 330
Ala Leu Lys Ile Trp Asp Leu Arg Gln Phe Lys Ser Gly Ser Pro
335 340 345
Val Ala Thr Phe Lys Gln His Val Ala Pro Val Thr Ser Val Glu
350 355 360
Trp His Pro Gln Asp Ser Gly Val Phe Ala Ala Ser Gly Ala Asp
365 370 375
His Gln Ile Thr Gln Trp Asp Leu Ala Val Glu Arg Asp Pro Glu
380 385 390
Ala Gly Asp Val Glu Ala Asp Pro Gly Leu Ala Asp Leu Pro Gln
395 400 405
Gln Leu Leu Phe Val His Gln Gly Glu Thr Glu Leu Lys Glu Leu
410 415 420
His Trp His Pro Gln Cys Pro Gly Leu Leu Val Ser Thr Ala Leu
425 430 435
Ser Gly Phe Thr Ile Phe Arg Thr Ile Ser Val
440 445

<210> 22
<211> 265
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 483862CD1

<400> 22
Met Ser Ser Gly Leu Arg Ala Ala Asp Phe Pro Arg Trp Lys Arg
1 5 10 15
His Ile Ser Glu Gln Leu Arg Arg Arg Asp Arg Leu Gln Arg Gln
20 25 30
Ala Phe Glu Glu Ile Ile Leu Gln Tyr Asn Lys Leu Leu Glu Lys
35 40 45
Ser Asp Leu His Ser Val Leu Ala Gln Lys Leu Gln Ala Glu Lys
50 55 60
His Asp Val Pro Asn Arg His Glu Ile Ser Pro Gly His Asp Gly
65 70 75
Thr Trp Asn Asp Asn Gln Leu Gln Glu Met Ala Gln Leu Arg Ile
80 85 90
Lys His Gln Glu Glu Leu Thr Glu Leu His Lys Lys Arg Gly Glu

```
                95                  100                 105
Leu Ala Gln Leu Val Ile Asp Leu Asn Asn Gln Met Gln Arg Lys
                110                 115                 120
Asp Arg Glu Met Gln Met Asn Glu Ala Lys Ile Ala Glu Cys Leu
                125                 130                 135
Gln Thr Ile Ser Asp Leu Glu Thr Glu Cys Leu Asp Leu Arg Thr
                140                 145                 150
Lys Leu Cys Asp Leu Glu Arg Ala Asn Gln Thr Leu Lys Asp Glu
                155                 160                 165
Tyr Asp Ala Leu Gln Ile Thr Phe Thr Ala Leu Glu Gly Lys Leu
                170                 175                 180
Arg Lys Thr Thr Glu Glu Asn Gln Glu Leu Val Thr Arg Trp Met
                185                 190                 195
Ala Glu Lys Ala Gln Glu Ala Asn Arg Leu Asn Ala Glu Asn Glu
                200                 205                 210
Lys Asp Ser Arg Arg Arg Gln Ala Arg Leu Gln Lys Glu Leu Ala
                215                 220                 225
Glu Ala Ala Lys Glu Pro Leu Pro Val Glu Gln Asp Asp Asp Ile
                230                 235                 240
Glu Val Ile Val Asp Glu Thr Ser Asp His Thr Glu Glu Thr Ser
                245                 250                 255
Pro Val Arg Ala Ile Ser Arg Ala Ala Thr
                260                 265
```

<210> 23
<211> 185
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1256777CD1

<400> 23

```
Met Leu Lys Ala Lys Ile Leu Phe Val Gly Pro Cys Glu Ser Gly
  1             5                   10                  15
Lys Thr Val Leu Ala Asn Phe Leu Thr Glu Ser Ser Asp Ile Thr
                20                  25                  30
Glu Tyr Ser Pro Thr Gln Gly Val Arg Ile Leu Glu Phe Glu Asn
                35                  40                  45
Pro His Val Thr Ser Asn Asn Lys Gly Thr Gly Cys Glu Phe Glu
                50                  55                  60
Leu Trp Asp Cys Gly Gly Asp Ala Lys Phe Glu Ser Cys Trp Pro
                65                  70                  75
Ala Leu Met Lys Asp Ala His Gly Val Val Ile Val Phe Asn Ala
                80                  85                  90
Asp Ile Pro Ser His Arg Lys Glu Met Glu Met Trp Tyr Ser Cys
                95                  100                 105
Phe Val Gln Gln Pro Ser Leu Gln Asp Thr Gln Cys Met Leu Ile
                110                 115                 120
Ala His His Lys Pro Gly Ser Gly Asp Asp Lys Gly Ser Leu Ser
                125                 130                 135
Leu Ser Pro Pro Leu Asn Lys Leu Lys Leu Val His Ser Asn Leu
                140                 145                 150
```

Glu Asp Asp Pro Glu Glu Ile Arg Met Glu Phe Ile Lys Tyr Leu
155 160 165
Lys Ser Ile Ile Asn Ser Met Ser Glu Ser Arg Asp Arg Glu Glu
170 175 180
Met Ser Ile Met Thr
185

<210> 24
<211> 554
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2198779CD1

<400> 24
Met Gly Ser Arg Asn Ser Ser Ser Ala Gly Ser Gly Ser Gly Asp
1 5 10 15
Pro Ser Glu Gly Leu Pro Arg Arg Gly Ala Gly Leu Arg Arg Ser
20 25 30
Glu Glu Glu Glu Glu Glu Asp Glu Asp Val Asp Leu Ala Gln Val
35 40 45
Leu Ala Tyr Leu Leu Arg Arg Gly Gln Val Arg Leu Val Gln Gly
50 55 60
Gly Gly Ala Ala Asn Leu Gln Phe Ile Gln Ala Leu Leu Asp Ser
65 70 75
Glu Glu Glu Asn Asp Arg Ala Trp Asp Gly Arg Leu Gly Asp Arg
80 85 90
Tyr Asn Pro Pro Val Asp Ala Thr Pro Asp Thr Arg Glu Leu Glu
95 100 105
Phe Asn Glu Ile Lys Thr Gln Val Glu Leu Ala Thr Gly Gln Leu
110 115 120
Gly Leu Arg Arg Ala Ala Gln Lys His Ser Phe Pro Arg Met Leu
125 130 135
His Gln Arg Glu Arg Gly Leu Cys His Arg Gly Ser Phe Ser Leu
140 145 150
Gly Glu Gln Ser Arg Val Ile Ser His Phe Leu Pro Asn Asp Leu
155 160 165
Gly Phe Thr Asp Ser Tyr Ser Gln Lys Ala Phe Cys Gly Ile Tyr
170 175 180
Ser Lys Asp Gly Gln Ile Phe Met Ser Ala Cys Gln Asp Gln Thr
185 190 195
Ile Arg Leu Tyr Asp Cys Arg Tyr Gly Arg Phe Arg Lys Phe Lys
200 205 210
Ser Ile Lys Ala Arg Asp Val Gly Trp Ser Val Leu Asp Val Ala
215 220 225
Phe Thr Pro Asp Gly Asn His Phe Leu Tyr Ser Ser Trp Ser Asp
230 235 240
Tyr Ile His Ile Cys Asn Ile Tyr Gly Glu Gly Asp Thr His Thr
245 250 255
Ala Leu Asp Leu Arg Pro Asp Glu Arg Arg Phe Ala Val Phe Ser
260 265 270
Ile Ala Val Ser Ser Asp Gly Arg Glu Val Leu Gly Gly Ala Asn
275 280 285

```
Asp Gly Cys Leu Tyr Val Phe Asp Arg Glu Gln Asn Arg Arg Thr
                290                 295                 300
Leu Gln Ile Glu Ser His Glu Asp Asp Val Asn Ala Val Ala Phe
                305                 310                 315
Ala Asp Ile Ser Ser Gln Ile Leu Phe Ser Gly Gly Asp Asp Ala
                320                 325                 330
Ile Cys Lys Val Trp Asp Arg Arg Thr Met Arg Glu Asp Asp Pro
                335                 340                 345
Lys Pro Val Gly Ala Leu Ala Gly His Gln Asp Gly Ile Thr Phe
                350                 355                 360
Ile Asp Ser Lys Gly Asp Ala Arg Tyr Leu Ile Ser Asn Ser Lys
                365                 370                 375
Asp Gln Thr Ile Lys Leu Trp Asp Ile Arg Arg Phe Ser Ser Arg
                380                 385                 390
Glu Gly Met Glu Ala Ser Arg Gln Ala Ala Thr Gln Gln Asn Trp
                395                 400                 405
Asp Tyr Arg Trp Gln Gln Val Pro Lys Lys Gly Phe Thr Leu His
                410                 415                 420
Pro Tyr Pro Ala Trp Arg Lys Leu Lys Leu Pro Gly Asp Ser Ser
                425                 430                 435
Leu Met Thr Tyr Arg Gly His Gly Val Leu His Thr Leu Ile Arg
                440                 445                 450
Cys Arg Phe Ser Pro Ile His Ser Thr Gly Gln Gln Phe Ile Tyr
                455                 460                 465
Ser Gly Cys Ser Thr Gly Lys Val Val Val Tyr Asp Leu Leu Ser
                470                 475                 480
Gly His Ile Val Lys Lys Leu Thr Asn His Lys Ala Cys Val Arg
                485                 490                 495
Asp Val Ser Trp His Pro Phe Glu Glu Lys Ile Val Ser Ser Ser
                500                 505                 510
Trp Asp Gly Asn Leu Arg Leu Trp Gln Tyr Arg Gln Ala Glu Tyr
                515                 520                 525
Phe Gln Asp Asp Met Pro Glu Ser Glu Glu Cys Ala Ser Ala Pro
                530                 535                 540
Ala Pro Val Pro Gln Ser Ser Thr Pro Phe Ser Ser Pro Gln
                545                 550


<210> 25
<211> 434
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No:  2226116CD1

<400> 25
Met Arg Pro Ser Ser Ser Val Ser Val Ser Cys Pro Ala Leu Asn
  1               5                  10                  15
Gln Val Ser His Phe Ala Asn Leu Thr Ser Val Gly Ala Met Ala
                 20                  25                  30
Pro Ala Arg Cys Phe Ser Ala Arg Leu Arg Thr Val Phe Gln Gly
                 35                  40                  45
Val Gly His Trp Ala Leu Ser Thr Trp Ala Gly Leu Lys Pro Ser
```

```
                50                  55                  60
Arg Leu Leu Pro Gln Arg Ala Ser Pro Arg Leu Leu Ser Val Gly
                65                  70                  75
Arg Ala Asp Leu Ala Lys His Gln Glu Leu Pro Gly Lys Lys Leu
                80                  85                  90
Leu Ser Glu Lys Lys Leu Lys Arg Tyr Phe Val Asp Tyr Arg Arg
                95                 100                 105
Val Leu Val Cys Gly Gly Asn Gly Gly Ala Gly Ala Ser Cys Phe
               110                 115                 120
His Ser Glu Pro Arg Lys Glu Phe Gly Gly Pro Asp Gly Gly Asp
               125                 130                 135
Gly Gly Asn Gly Gly His Val Ile Leu Arg Val Asp Gln Gln Val
               140                 145                 150
Lys Ser Leu Ser Ser Val Leu Ser Arg Tyr Gln Gly Phe Ser Gly
               155                 160                 165
Glu Asp Gly Gly Ser Lys Asn Cys Phe Gly Arg Ser Gly Ala Val
               170                 175                 180
Leu Tyr Ile Arg Val Pro Val Gly Thr Leu Val Lys Glu Gly Gly
               185                 190                 195
Arg Val Val Ala Asp Leu Ser Cys Val Gly Asp Glu Tyr Ile Ala
               200                 205                 210
Ala Leu Gly Gly Ala Gly Gly Lys Gly Asn Arg Phe Phe Leu Ala
               215                 220                 225
Asn Asn Asn Arg Ala Pro Val Thr Cys Thr Pro Gly Gln Pro Gly
               230                 235                 240
Gln Gln Arg Val Leu His Leu Glu Leu Lys Thr Val Ala His Ala
               245                 250                 255
Gly Met Val Gly Phe Pro Asn Ala Gly Lys Ser Ser Leu Leu Arg
               260                 265                 270
Ala Ile Ser Asn Ala Arg Pro Ala Val Ala Ser Tyr Pro Phe Thr
               275                 280                 285
Thr Leu Lys Pro His Val Gly Ile Val His Tyr Glu Gly His Leu
               290                 295                 300
Gln Ile Ala Val Ala Asp Ile Pro Gly Ile Ile Arg Gly Ala His
               305                 310                 315
Gln Asn Arg Gly Leu Gly Ser Ala Phe Leu Arg His Ile Glu Arg
               320                 325                 330
Cys Arg Phe Leu Leu Phe Val Val Asp Leu Ser Gln Pro Glu Pro
               335                 340                 345
Trp Thr Gln Val Asp Asp Leu Lys Tyr Glu Leu Glu Met Tyr Glu
               350                 355                 360
Lys Gly Leu Ser Ala Arg Pro His Ala Ile Val Ala Asn Lys Ile
               365                 370                 375
Asp Leu Pro Glu Ala Gln Ala Asn Leu Ser Gln Leu Arg Asp His
               380                 385                 390
Leu Gly Gln Glu Val Ile Val Leu Ser Ala Leu Thr Gly Glu Asn
               395                 400                 405
Leu Glu Gln Leu Leu Leu His Leu Lys Val Leu Tyr Asp Ala Tyr
               410                 415                 420
Ala Glu Ala Glu Leu Gly Gln Gly Arg Gln Pro Leu Arg Trp
               425                 430
```

<210> 26
<211> 826

<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2504472CD1

<400> 26
Met Val Ala Pro Val Leu Glu Thr Ser His Val Phe Cys Cys Pro
1 5 10 15
Asn Arg Val Arg Gly Val Leu Asn Trp Ser Ser Gly Pro Arg Gly
20 25 30
Leu Leu Ala Phe Gly Thr Ser Cys Ser Val Val Leu Tyr Asp Pro
35 40 45
Leu Lys Arg Val Val Val Thr Asn Leu Asn Gly His Thr Ala Arg
50 55 60
Val Asn Cys Ile Gln Trp Ile Cys Lys Gln Asp Gly Ser Pro Ser
65 70 75
Thr Glu Leu Val Ser Gly Gly Ser Asp Asn Gln Val Ile His Trp
80 85 90
Glu Ile Glu Asp Asn Gln Leu Leu Lys Ala Val His Leu Gln Gly
95 100 105
His Glu Gly Pro Val Tyr Ala Val His Ala Val Tyr Gln Arg Arg
110 115 120
Thr Ser Asp Pro Ala Leu Cys Thr Leu Ile Val Ser Ala Ala Ala
125 130 135
Asp Ser Ala Val Arg Leu Trp Ser Lys Lys Gly Pro Glu Val Met
140 145 150
Cys Leu Gln Thr Leu Asn Phe Gly Asn Gly Phe Ala Leu Ala Leu
155 160 165
Cys Leu Ser Phe Leu Pro Asn Thr Asp Val Pro Ile Leu Ala Cys
170 175 180
Gly Asn Asp Asp Cys Arg Ile His Ile Phe Ala Gln Gln Asn Asp
185 190 195
Gln Phe Gln Lys Val Leu Ser Leu Cys Gly His Glu Asp Trp Ile
200 205 210
Arg Gly Val Glu Trp Ala Ala Phe Gly Arg Asp Leu Phe Leu Ala
215 220 225
Ser Cys Ser Gln Asp Cys Leu Ile Arg Ile Trp Lys Leu Tyr Ile
230 235 240
Lys Ser Thr Ser Leu Glu Thr Gln Asp Asp Asp Asn Ile Arg Leu
245 250 255
Lys Glu Asn Thr Phe Thr Ile Glu Asn Glu Ser Val Lys Ile Ala
260 265 270
Phe Ala Val Thr Leu Glu Thr Val Leu Ala Gly His Glu Asn Trp
275 280 285
Val Asn Ala Val His Trp Gln Pro Val Phe Tyr Lys Asp Gly Val
290 295 300
Leu Gln Gln Pro Val Arg Leu Leu Ser Ala Ser Met Asp Lys Thr
305 310 315
Met Ile Leu Trp Ala Pro Asp Glu Glu Ser Gly Val Trp Leu Glu
320 325 330
Gln Val Arg Val Gly Glu Val Gly Gly Asn Thr Leu Gly Phe Tyr
335 340 345
Asp Cys Gln Phe Asn Glu Asp Gly Ser Met Ile Ile Ala His Ala

350 355 360

Phe His Gly Ala Leu His Leu Trp Lys Gln Asn Thr Val Asn Pro

365 370 375

Arg Glu Trp Thr Pro Glu Ile Val Ile Ser Gly His Phe Asp Gly

380 385 390

Val Gln Asp Leu Val Trp Asp Pro Glu Gly Glu Phe Ile Ile Thr

395 400 405

Val Gly Thr Asp Gln Thr Thr Arg Leu Phe Ala Pro Trp Lys Arg

410 415 420

Lys Asp Gln Ser Gln Val Thr Trp His Glu Ile Ala Arg Pro Gln

425 430 435

Ile His Gly Tyr Asp Leu Lys Cys Leu Ala Met Ile Asn Arg Phe

440 445 450

Gln Phe Val Ser Gly Ala Asp Glu Lys Val Leu Arg Val Phe Ser

455 460 465

Ala Pro Arg Asn Phe Val Glu Asn Phe Cys Ala Ile Thr Gly Gln

470 475 480

Ser Leu Asn His Val Leu Cys Asn Gln Asp Ser Asp Leu Pro Glu

485 490 495

Gly Ala Thr Val Pro Ala Leu Gly Leu Ser Asn Lys Ala Val Phe

500 505 510

Gln Gly Asp Ile Ala Ser Gln Pro Ser Asp Glu Glu Glu Leu Leu

515 520 525

Thr Ser Thr Gly Phe Glu Tyr Gln Gln Val Ala Phe Gln Pro Ser

530 535 540

Ile Leu Thr Glu Pro Pro Thr Glu Asp His Leu Leu Gln Asn Thr

545 550 555

Leu Trp Pro Glu Val Gln Lys Leu Tyr Gly His Gly Tyr Glu Ile

560 565 570

Phe Cys Val Thr Cys Asn Ser Ser Lys Thr Leu Leu Ala Ser Ala

575 580 585

Cys Lys Ala Ala Lys Lys Glu His Ala Ala Ile Ile Leu Trp Asn

590 595 600

Thr Thr Ser Trp Lys Gln Val Gln Asn Leu Val Phe His Ser Leu

605 610 615

Thr Val Thr Gln Met Ala Phe Ser Pro Asn Glu Lys Phe Leu Leu

620 625 630

Ala Val Ser Arg Asp Arg Thr Trp Ser Leu Trp Lys Lys Gln Asp

635 640 645

Thr Ile Ser Pro Glu Phe Glu Pro Val Phe Ser Leu Phe Ala Phe

650 655 660

Thr Asn Lys Ile Thr Ser Val His Ser Arg Ile Ile Trp Ser Cys

665 670 675

Asp Trp Ser Pro Asp Ser Lys Tyr Phe Phe Thr Gly Ser Arg Asp

680 685 690

Lys Lys Val Val Val Trp Gly Glu Cys Asp Ser Thr Asp Asp Cys

695 700 705

Ile Glu His Asn Ile Gly Pro Cys Ser Ser Val Leu Asp Val Gly

710 715 720

Gly Ala Val Thr Ala Val Ser Val Cys Pro Val Leu His Pro Ser

725 730 735

Gln Arg Tyr Val Val Ala Val Gly Leu Glu Cys Gly Lys Ile Cys

740 745 750

Leu Tyr Thr Trp Lys Lys Thr Asp Gln Val Pro Glu Ile Asn Asp

755 760 765

Trp Thr His Cys Val Glu Thr Ser Gln Ser Gln Ser His Thr Leu
770 775 780
Ala Ile Arg Lys Leu Cys Trp Lys Asn Cys Ser Gly Lys Thr Glu
785 790 795
Gln Lys Glu Ala Glu Gly Ala Glu Trp Leu His Phe Ala Ser Cys
800 805 810
Gly Glu Asp His Thr Val Lys Ile His Arg Val Asn Lys Cys Ala
815 820 825
Leu

<210> 27
<211> 618
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3029920CD1

<400> 27
Met Lys Lys Asp Val Arg Ile Leu Leu Val Gly Glu Pro Arg Val
1 5 10 15
Gly Lys Thr Ser Leu Ile Met Ser Leu Val Ser Glu Glu Phe Pro
20 25 30
Glu Glu Val Pro Pro Arg Ala Glu Glu Ile Thr Ile Pro Ala Asp
35 40 45
Val Thr Pro Glu Arg Val Pro Thr His Ile Val Asp Tyr Ser Glu
50 55 60
Ala Glu Gln Ser Asp Glu Gln Leu His Gln Glu Ile Ser Gln Ala
65 70 75
Asn Val Ile Cys Ile Val Tyr Ala Val Asn Asn Lys His Ser Ile
80 85 90
Asp Lys Val Thr Ser Arg Trp Ile Pro Leu Ile Asn Glu Arg Thr
95 100 105
Asp Lys Asp Ser Arg Leu Pro Leu Ile Leu Val Gly Asn Lys Ser
110 115 120
Asp Leu Val Glu Tyr Ser Ser Met Glu Thr Ile Leu Pro Ile Met
125 130 135
Asn Gln Tyr Thr Glu Ile Glu Thr Cys Val Glu Cys Ser Ala Lys
140 145 150
Asn Leu Lys Asn Ile Ser Glu Leu Phe Tyr Tyr Ala Gln Lys Ala
155 160 165
Val Leu His Pro Thr Gly Pro Leu Tyr Cys Pro Glu Glu Lys Glu
170 175 180
Met Lys Pro Ala Cys Ile Lys Ala Leu Thr Arg Ile Phe Lys Ile
185 190 195
Ser Asp Gln Asp Asn Asp Gly Thr Leu Asn Asp Ala Glu Leu Asn
200 205 210
Phe Phe Gln Arg Ile Cys Phe Asn Thr Pro Leu Ala Pro Gln Ala
215 220 225
Leu Glu Asp Val Lys Asn Val Val Arg Lys His Ile Ser Asp Gly
230 235 240
Val Ala Asp Ser Gly Leu Thr Leu Lys Gly Phe Leu Phe Leu His
245 250 255

```
Thr Leu Phe Ile Gln Arg Gly Arg His Glu Thr Thr Trp Thr Val
                260                 265                 270
Leu Arg Arg Phe Gly Tyr Asp Asp Asp Leu Asp Leu Thr Pro Glu
                275                 280                 285
Tyr Leu Phe Pro Leu Leu Lys Ile Pro Pro Asp Cys Thr Thr Glu
                290                 295                 300
Leu Asn His His Ala Tyr Leu Phe Leu Gln Ser Thr Phe Asp Lys
                305                 310                 315
His Asp Leu Asp Arg Asp Cys Ala Leu Ser Pro Asp Glu Leu Lys
                320                 325                 330
Asp Leu Phe Lys Val Phe Pro Tyr Ile Pro Trp Gly Pro Asp Val
                335                 340                 345
Asn Asn Thr Val Cys Thr Asn Glu Arg Gly Trp Ile Thr Tyr Gln
                350                 355                 360
Gly Phe Leu Ser Gln Trp Thr Leu Thr Thr Tyr Leu Asp Val Gln
                365                 370                 375
Arg Cys Leu Glu Tyr Leu Gly Tyr Leu Gly Tyr Ser Ile Leu Thr
                380                 385                 390
Glu Gln Glu Ser Gln Ala Ser Ala Val Thr Val Thr Arg Asp Lys
                395                 400                 405
Lys Ile Asp Leu Gln Lys Lys Gln Thr Gln Arg Asn Val Phe Arg
                410                 415                 420
Cys Asn Val Ile Gly Val Lys Asn Cys Gly Lys Ser Gly Val Leu
                425                 430                 435
Gln Ala Leu Leu Gly Arg Asn Leu Met Arg Gln Lys Lys Ile Arg
                440                 445                 450
Glu Asp His Lys Ser Tyr Tyr Ala Ile Asn Thr Val Tyr Val Tyr
                455                 460                 465
Gly Gln Glu Lys Tyr Leu Leu Leu His Asp Ile Ser Glu Ser Glu
                470                 475                 480
Phe Leu Thr Glu Ala Glu Ile Ile Cys Asp Val Val Cys Leu Val
                485                 490                 495
Tyr Asp Val Ser Asn Pro Lys Ser Phe Glu Tyr Cys Ala Arg Ile
                500                 505                 510
Phe Lys Gln His Phe Met Asp Ser Arg Ile Pro Cys Leu Ile Val
                515                 520                 525
Ala Ala Lys Ser Asp Leu His Glu Val Lys Gln Glu Tyr Ser Ile
                530                 535                 540
Ser Pro Thr Asp Phe Cys Arg Lys His Lys Met Pro Pro Pro Gln
                545                 550                 555
Ala Phe Thr Cys Asn Thr Ala Asp Ala Pro Ser Lys Asp Ile Phe
                560                 565                 570
Val Lys Leu Thr Thr Met Ala Met Tyr Pro His Val Thr Gln Ala
                575                 580                 585
Asp Leu Lys Ser Ser Thr Phe Trp Leu Arg Ala Ser Phe Gly Ala
                590                 595                 600
Thr Val Phe Ala Val Leu Gly Phe Ala Met Tyr Lys Ala Leu Leu
                605                 610                 615
Lys Gln Arg
```

<210> 28
<211> 596
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3332415CD1

<400> 28
Met Glu Pro Glu Leu Asp Ala Gln Lys Gln Pro Arg Pro Arg Arg
1 5 10 15
Arg Ser Arg Arg Ala Ser Gly Leu Ser Thr Glu Gly Ala Thr Gly
20 25 30
Pro Ser Ala Asp Thr Ser Gly Ser Glu Leu Asp Gly Arg Cys Ser
35 40 45
Leu Arg Arg Gly Ser Ser Phe Thr Phe Leu Thr Pro Gly Pro Asn
50 55 60
Trp Asp Phe Thr Leu Lys Arg Lys Arg Arg Glu Lys Asp Asp Asp
65 70 75
Val Val Ser Leu Ser Ser Leu Asp Leu Lys Glu Pro Ser Asn Lys
80 85 90
Arg Val Arg Pro Leu Ala Arg Val Thr Ser Leu Ala Asn Leu Ile
95 100 105
Ser Pro Val Arg Asn Gly Ala Val Arg Arg Phe Gly Gln Thr Ile
110 115 120
Gln Ser Phe Thr Leu Arg Gly Asp His Arg Ser Pro Ala Ser Ala
125 130 135
Gln Lys Phe Ser Ser Arg Ser Thr Val Pro Thr Pro Ala Lys Arg
140 145 150
Arg Ser Ser Ala Leu Trp Ser Glu Met Leu Asp Ile Thr Met Lys
155 160 165
Glu Ser Leu Thr Thr Arg Glu Ile Arg Arg Gln Glu Ala Ile Tyr
170 175 180
Glu Met Ser Arg Gly Glu Gln Asp Leu Ile Glu Asp Leu Lys Leu
185 190 195
Ala Arg Lys Ala Tyr His Asp Pro Met Leu Lys Leu Ser Ile Met
200 205 210
Ser Glu Glu Glu Leu Thr His Ile Phe Gly Asp Leu Asp Ser Tyr
215 220 225
Ile Pro Leu His Glu Asp Leu Leu Thr Arg Ile Gly Glu Ala Thr
230 235 240
Lys Pro Asp Gly Thr Val Glu Gln Ile Gly His Ile Leu Val Ser
245 250 255
Trp Leu Pro Arg Leu Asn Ala Tyr Arg Gly Tyr Cys Ser Asn Gln
260 265 270
Leu Ala Ala Lys Ala Leu Leu Asp Gln Lys Lys Gln Asp Pro Arg
275 280 285
Val Gln Asp Phe Leu Gln Arg Cys Leu Glu Ser Pro Phe Ser Arg
290 295 300
Lys Leu Asp Leu Trp Ser Phe Leu Asp Ile Pro Arg Ser Arg Leu
305 310 315
Val Lys Tyr Pro Leu Leu Leu Lys Glu Ile Leu Lys His Thr Pro
320 325 330
Lys Glu His Pro Asp Val Gln Leu Leu Glu Asp Ala Ile Leu Ile
335 340 345
Ile Gln Gly Val Leu Ser Asp Ile Asn Leu Lys Lys Gly Glu Ser
350 355 360
Glu Cys Gln Tyr Tyr Ile Asp Lys Leu Glu Tyr Leu Asp Glu Lys
365 370 375

```
Gln Arg Asp Pro Arg Ile Glu Ala Ser Lys Val Leu Leu Cys His
                380                 385                 390
Gly Glu Leu Arg Ser Lys Ser Gly His Lys Leu Tyr Ile Phe Leu
                395                 400                 405
Phe Gln Asp Ile Leu Val Leu Thr Arg Pro Val Thr Arg Asn Glu
                410                 415                 420
Arg His Ser Tyr Gln Val Tyr Arg Gln Pro Ile Pro Val Gln Glu
                425                 430                 435
Leu Val Leu Glu Asp Leu Gln Asp Gly Asp Val Arg Met Gly Gly
                440                 445                 450
Ser Phe Arg Gly Ala Phe Ser Asn Ser Glu Lys Ala Lys Asn Ile
                455                 460                 465
Phe Arg Ile Arg Phe His Asp Pro Ser Pro Ala Gln Ser His Thr
                470                 475                 480
Leu Gln Ala Asn Asp Val Phe His Lys Gln Gln Trp Phe Asn Cys
                485                 490                 495
Ile Arg Ala Ala Ile Ala Pro Phe Gln Ser Ala Gly Ser Pro Pro
                500                 505                 510
Glu Leu Gln Gly Leu Pro Glu Leu His Glu Glu Cys Glu Gly Asn
                515                 520                 525
His Pro Ser Ala Arg Lys Leu Thr Ala Gln Arg Arg Ala Ser Thr
                530                 535                 540
Val Ser Ser Val Thr Gln Val Glu Val Asp Glu Asn Ala Tyr Arg
                545                 550                 555
Cys Gly Ser Gly Met Gln Met Ala Glu Asp Ser Lys Ser Leu Lys
                560                 565                 570
Thr His Gln Thr Gln Pro Gly Ile Arg Arg Ala Arg Asp Lys Ala
                575                 580                 585
Leu Ser Gly Gly Lys Arg Lys Glu Thr Leu Val
                590                 595
```

<210> 29
<211> 589
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4031536CD1

<400> 29

```
Met Ser Lys Pro Gly Lys Pro Thr Leu Asn His Gly Leu Val Pro
  1               5                  10                  15
Val Asp Leu Lys Ser Ala Lys Glu Pro Leu Pro His Gln Thr Val
                 20                  25                  30
Met Arg Ile Phe Ser Ile Ser Ile Ile Ala Gln Gly Leu Pro Phe
                 35                  40                  45
Cys Arg Arg Arg Met Lys Arg Lys Leu Asp His Gly Ser Glu Val
                 50                  55                  60
Arg Ser Phe Ser Leu Gly Lys Lys Pro Cys Lys Val Ser Glu Tyr
                 65                  70                  75
Thr Ser Thr Thr Gly Leu Val Pro Cys Ser Ala Thr Pro Thr Thr
                 80                  85                  90
Phe Gly Asp Leu Arg Ala Ala Asn Gly Gln Gly Gln Gln Arg Arg
```

```
                95                  100                 105
Arg Ile Thr Ser Val Gln Pro Pro Thr Gly Leu Gln Glu Trp Leu
                110                 115                 120
Lys Met Phe Gln Ser Trp Ser Gly Pro Glu Lys Leu Leu Ala Leu
                125                 130                 135
Asp Glu Leu Ile Asp Ser Cys Glu Pro Thr Gln Val Lys His Met
                140                 145                 150
Met Gln Val Ile Glu Pro Gln Phe Gln Arg Asp Phe Ile Ser Leu
                155                 160                 165
Leu Pro Lys Glu Leu Ala Leu Tyr Val Leu Ser Phe Leu Glu Pro
                170                 175                 180
Lys Asp Leu Leu Gln Ala Ala Gln Thr Cys Arg Tyr Trp Arg Ile
                185                 190                 195
Leu Ala Glu Asp Asn Leu Leu Trp Arg Glu Lys Cys Lys Glu Glu
                200                 205                 210
Gly Ile Asp Glu Pro Leu His Ile Lys Arg Arg Lys Val Ile Lys
                215                 220                 225
Pro Gly Phe Ile His Ser Pro Trp Lys Ser Ala Tyr Ile Arg Gln
                230                 235                 240
His Arg Ile Asp Thr Asn Trp Arg Arg Gly Glu Leu Lys Ser Pro
                245                 250                 255
Lys Val Leu Lys Gly His Asp Asp His Val Ile Thr Cys Leu Gln
                260                 265                 270
Phe Cys Gly Asn Arg Ile Val Ser Gly Ser Asp Asp Asn Thr Leu
                275                 280                 285
Lys Val Trp Ser Ala Val Thr Gly Lys Cys Leu Arg Thr Leu Val
                290                 295                 300
Gly His Thr Gly Gly Val Trp Ser Ser Gln Met Arg Asp Asn Ile
                305                 310                 315
Ile Ile Ser Gly Ser Thr Asp Arg Thr Leu Lys Val Trp Asn Ala
                320                 325                 330
Glu Thr Gly Glu Cys Ile His Thr Leu Tyr Gly His Thr Ser Thr
                335                 340                 345
Val Arg Cys Met His Leu His Glu Lys Arg Val Val Ser Gly Ser
                350                 355                 360
Arg Asp Ala Thr Leu Arg Val Trp Asp Ile Glu Thr Gly Gln Cys
                365                 370                 375
Leu His Val Leu Met Gly His Val Ala Ala Val Arg Cys Val Gln
                380                 385                 390
Tyr Asp Gly Arg Arg Val Val Ser Gly Ala Tyr Asp Phe Met Val
                395                 400                 405
Lys Val Trp Asp Pro Glu Thr Glu Thr Cys Leu His Thr Leu Gln
                410                 415                 420
Gly His Thr Asn Arg Val Tyr Ser Leu Gln Phe Asp Gly Ile His
                425                 430                 435
Val Val Ser Gly Ser Leu Asp Thr Ser Ile Arg Val Trp Asp Val
                440                 445                 450
Glu Thr Gly Asn Cys Ile His Thr Leu Thr Gly His Gln Ser Leu
                455                 460                 465
Thr Ser Gly Met Glu Leu Lys Asp Asn Ile Leu Val Ser Gly Asn
                470                 475                 480
Ala Asp Ser Thr Val Lys Ile Trp Asp Ile Lys Thr Gly Gln Cys
                485                 490                 495
Leu Gln Thr Leu Gln Gly Pro Asn Lys His Gln Ser Ala Val Thr
                500                 505                 510
```

Cys Leu Gln Phe Asn Lys Asn Phe Val Ile Thr Ser Ser Asp Asp
515 520 525
Gly Thr Val Lys Leu Trp Asp Leu Lys Thr Gly Glu Phe Ile Arg
530 535 540
Asn Leu Val Thr Leu Glu Ser Gly Gly Ser Gly Gly Val Val Trp
545 550 555
Arg Ile Arg Ala Ser Asn Thr Lys Leu Val Cys Ala Val Gly Ser
560 565 570
Arg Asn Gly Thr Glu Glu Thr Lys Leu Leu Val Leu Asp Phe Asp
575 580 585
Val Asp Met Lys

<210> 30
<211> 3375
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 708398CB1

<400> 30
ggggagaagg gagcccgcag gatcaggggt cagagttagg gggcttccct cctgctgcac 60
cctcatctca gggccgccaa cttccagctg cagcggcgac tttcagtttc atttccacgg 120
accctcctgc ctgggcccgc agccgccgcc gcgatgccca gtaagttcag ctgccggcag 180
ctccgggagg cgggccagtg tttcgagagt ttcctggtcg ttcggggact ggacatggag 240
acagatcgcg agcggctgcg gaccatttat aaccgcgact tcaagatcag ctttgggacc 300
cccgcccctg gcttctcctc catgctgtat ggaatgaaga ttgcaaatct ggcctacgtc 360
accaagactc gggtcaggtt cttcagactc gaccgctggg ccgacgtgcg gttcccagaa 420
aagaggagaa tgaagctggg gtcagatatc agcaaacacc acaagtcact gctagccaag 480
atcttttatg acagggctga gtatcttcat gggaaacatg gtgtggatgt ggaagtccag 540
gggccccatg aagcccgaga tgggcagctc cttatccgcc tggatttgaa ccgcaaagag 600
gtgctgaccc tgaggcttcg gaatggcgga acccagtctg ttaccctcac tcacctcttc 660
ccactctgcc ggacacccca gtttgctttc tacaatgaag accaggagtt gccctgtcca 720
ctgggccccg gtgaatgcta tgaactccat gtccattgta agaccagctt tgtgggctac 780
ttcccagcca cagtgctctg ggagctgctg ggacctgggg agtcgggttc agaaggagcc 840
ggcacattct acattgcccg ctccttggct gccgtcgccc acagcccccct ggctgcacag 900
ctgaagccca tgactccctt caagcggacc cggatcaccg gaaaccctgt ggtgaccaat 960
cggatagagg aaggagagag acctgaccgc gctaagggct atgacctgga gttaagtatg 1020
gcgctgggga catactaccc acctccccgc ctcaggcagc tgctccccat gcttcttcag 1080
ggaacaagta tcttcactgc ccctaaggag atcgcagaga tcaaggccca gctggagaca 1140
gccctgaagt ggaggaacta tgaggtgaag ctgcggctgc tgctgcacct ggaggaactg 1200
cagatggagc atgatatccg gcactatgac ctggagtcgg tgcccatgac ctgggacccct 1260
gtggaccaga accccaggct gctcacgctg gaggttcctg gagtgactga gagccgcccc 1320
tcagtgctac ggggcgacca cctgtttgcc cttttgtcct cggagacaca ccaggaggac 1380
cccatcacat ataagggctt tgtgcacaag gtggaattgg accgtgtcaa gctgagcttt 1440
tccatgagcc tcctgagccg ctttgtggat gggctgacct tcaaggtgaa ctttaccttc 1500
aaccgccagc cgctgcgagt ccagcaccgt gccctggagc tgacagggcg ctggctgctg 1560
tggcccatgc tctttcctgt ggcacctcgg gacgtcccgc tgctgccctc agatgtgaaa 1620
ctcaagctgt acgaccggag tctggagtca aacccagagc agctgcaggc catgaggcac 1680
attgttacgg gcaccacccg tccagccccc tacatcatct ttgggcctcc aggcaccggc 1740
aagactgtca cgttagtgga ggcaattaag caggtggtga agcacttgcc caaagcccac 1800
atcttggcct gcgctccatc caactcaggg gctgacctac tctgtcaaag gctccgggtc 1860
caccttccta gctccatcta ccgcctcctg gcccccagca gggacatccg catggtacct 1920

```
gaggacatca agccctgctg caactgggac gcaaagaagg gggagtatgt atttcccgcc 1980
aagaagaagc tgcaggaata ccgggtctta attaccaccc tcatcactgc cggcaggttg 2040
gtctcggccc agtttcccat tgatcacttc acacacatct tcatcgatga ggctggccac 2100
tgcatggagc ctgagagtct ggtagctata gcagggctga tggaagtaaa ggaaacaggt 2160
gatccaggag ggcagctggt gctggcagga gaccctcggc agctggggcc tgtgctgcgt 2220
tccccactga cccagaagca tggactggga tactcactgc tggagcggct gctcatctac 2280
aactccctgt acaagaaggg ccctgatggc tatgaccccc agttcataac caagctgctc 2340
cgcaactaca ggtctcatcc caccatcctg gacattccta accagctcta ttatgaaggg 2400
gagctgcagg cctgtgctga tgtcgtggat cgagaacgct tctgccgctg ggcgggccta 2460
cctcgacagg gctttcccat catctttcac ggcgtaatgg gcaaagatga gcgtgaaggc 2520
aacagcccat ccttcttcaa ccctgaagag gctgccacag tgacttccta cctgaagctg 2580
ctcctggccc cctcctccaa gaagggcaaa gctcgcctga gccctcgaag tgtgggcgtc 2640
atctccccgt accggaaaca ggtggagaaa atccgttact gcatcaccaa acttgacagg 2700
gagcttcgag gactggatga catcaaggac ttgaaggtgg gttcagtaga agaattccaa 2760
ggccaagaac gaagcgtcat cctcatctcc accgtgcgaa gcagccagag ctttgtgcag 2820
ctggatctgg actttaatct gggtttcctt aagaacccca agaggttcaa tgtagctgtg 2880
acccgggcca aggccctgct catcatcgtg gggaaccccc ttctcctggg ccatgaccct 2940
gactggaaag tattcctgga gttctgtaaa gaaaacggag ggtataccgg gtgtcccttc 3000
cctgccaaac tggacctgca acagggacag aatttactgc aaggtctgag caagctcagc 3060
ccctctacct cagggcccca cagccatgac tacctccccc aggagcggga gggtgaaggg 3120
ggcctgtctc tgcaagtgga gccagagtgg aggaatgagc tctgaagaca cagcacccag 3180
ccttctcgca ccagccaagc cttaactgcc tgcctgaccc tgaaccagaa cccagctgaa 3240
ctgcccctcc aagggacagg aaggctgggg gagggagttt acaacccaag ccattccacc 3300
ccctcccctg ctggggagaa tgacacatca agctgctaac aattggggga aggggaagga 3360
agaaaactct gaaac                                                  3375
```

<210> 31
<211> 2434
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1259937CB1

<400> 31

```
caaggatccg atgggtatat ggagtgtgag gtaatggatc attcatgtgg aaggatgcag   60
ggggtttttg agaccagggt ttggaagaga gttcagcact gctggtagtt ttgggaatca 120
cccatgtgca ggcgacacat gaggcagtaa ggaactctgc aggggtccct gagatttgga 180
aatgtaggga agagcaatgg attgaggtcc gaacctggag gatctgctat acgcagagct 240
gggaggaggg acagagtcag taccagagtc ggaaaaaagc agggtgggaa ggggaacctg 300
agtcaggaga cttgcctggc aggcgctgcc ctgccagcag aggcctgaca gtggtttcca 360
tgaactgcat ccctgctgtg ggctgggaca gggccactga cacagtatcg gagcacagaa 420
ggggaaagga gcaggaggga ttccaactct gccagttagc agctgtgtgg ctttgggcat 480
gttacttaac ctctctgagc ctcatttatt tcatccataa aatggaaata aaaataatac 540
ttttgtcaaa ggcgcattgt gaatatttag atcctcagaa taatgcctgg cttgtagcaa 600
atggtagctg gaggaaaagg aagagaaaac caaagtcagc agctgaagga ttttcatatt 660
agaactgctc tggacctatc tggcagatgc agaagcacac acacacggag gggcatggat 720
ttgccccgcc cttagacatg ttgtgtcttc tcctggatcc ttggtcccag gtgccttacc 780
tgagctcagg tgaatgtggc aagcagagcc ctctggtggt gtgaatgctg tgtggcgccc 840
gtgctcctgg tgacacaggg acctcacaat ccctccctcc acggtctcct ctcatgtcct 900
cccagcctta ttttctcgtt cctcttcctc ccaggcccgg aacttgcctg tttggctccc 960
caaccaggac gagcccctte ctggcagcag ctgtgccatc caagttgggg ataaagtccc 1020
ctatgacatc tgccggccag accactcagt gttgaccctg cagctgcctg tgacagcctc 1080
```

```
cgtgagagag gtgatggcag cgttggccca ggaggatggc tggaccaagg ggcaggtgct 1140
ggtgaaggtc aattctgcag gtgatgccat tggcctgcag ccagatgccc gtggtgtggc 1200
cacatctctg gggctcaatg agcgtctctt tgttgtcaac ccacaggaag tgcatgagct 1260
gatcccacac cctgaccagc tggggcccac tgtgggctct gctgaggggc tggacctggt 1320
gagtgccaag gacctggcag gccagctgac ggaccacgac tggagcctct tcaacagtat 1380
ccaccaggtg gagctgatcc actatgtgct gggcccccag catctgcggg atgtcaccac 1440
cgccaacctg gagcgcttca tgcgccgctt caatgagctg cagtactggg tggccaccga 1500
gctgtgtctc tgccccgtgc ccggcccccg ggcccagctg ctcaggaagt tcattaagct 1560
ggcggcccac ctcaaggagc agaagaatct caattccttc tttgccgtca tgtttggcct 1620
cagcaactcg gccatcagcc gcctagccca cacctgggag cggctgcctc acaaagtccg 1680
gaagctgtac tccgccctcg agaggctgct ggatccctca tggaaccacc gggtataccg 1740
actggccctc gccaagctct cccctcctgt catccccttc atgccccttc ttctcaaaga 1800
catgaccttc attcatgagg gaaaccacac actagtggag aatctcatca actttgagaa 1860
gatgagaatg atggccagag ccgcgcggat gctgcaccac tgccgaagcc acaaccctgt 1920
gcctctctca ccactcagaa gccgagtttc ccacctccac gaggacagcc aggtggcgag 1980
gatttccaca tgctcggagc agtccctgag cacccggagt ccagccagca cctgggctta 2040
tgtccagcag ctgaaggtca ttgacaacca gcgggaactc tcccgcctct cccgagagct 2100
ggagccatga ggaggggctg ggactggagc tggagcaggc acttgcagcc gggaaagcca 2160
gggtgtgccg ggccaagata ctcacaggct ggccacagct gggcaaggct ctccgtggag 2220
tggactcgag tccctggagc aggcagtgtg gaggcagcca tcccctgtga tgactggcag 2280
ctaaggagga cctcggagtg gacccgagcc aggaataacg aatgacccaa ggccaaggaa 2340
gggaggacag agaggcccca ggagtgggtg gagagtggag tgcgctggga cgttgtgtgc 2400
aatagagagg tctccacacc agaaaaaaaa aaaa                             2434
```

<210> 32
<211> 892
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1452285CB1

<400> 32

```
cacgcctctc tctgcaccta cacttgcgct ccccaagtct ctctcgtgcg cagagcccag  60
gctgcgcttc cctggtcagg cacggcacgt ctggccggcc gccaggatgc aggccccgca 120
caaggagcac ctgtacaagt tgctggtgat tggcgacctg ggcgtgggga agaccagtat 180
catcaagcgc tacgtgcacc agaacttctc ctcgcactac cgggccacaa tcggcgtgga 240
cttcgcgctc aaggtgctcc actgggaccc ggagactgtg gtgcgcctgc agctctggga 300
tatcgcaggt caagaaagat ttggaaacat gacgagggtc tattaccgag aagctatggg 360
tgcatttatt gtcttcgatg tcaccaggcc agccacattt gaagcagtgg caaagtggaa 420
aaatgatttg gactccaagt taagtctccc taatggcaaa ccggtttcag tggttttgtt 480
ggccaacaaa tgtgaccagg ggaaggatgt gctcatgaac aatggcctca agatggacca 540
gttctgcaag gagcacggtt tcgtaggatg gtttgaaaca tcagcaaagg aaaatataaa 600
cattgatgaa gcctccagat gcctggtgaa acacatactt gcaaatgagt gtgacctaat 660
ggagtctatt gagccggacg tcgtgaagcc ccatctcaca tcaaccaagg ttgccagctg 720
ctctggctgt gccaaatcct agtaggcacc tttgctggtg tctggtagga atgacctcat 780
tgttccacaa attgtgcctc tatttttacc attttgggta aacgtcagga tagagatacc 840
acatgtggca agccaaagat ctatgcctcc atatgtgcct ttctgttagc tt         892
```

<210> 33

<211> 2288
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1812894CB1

<400> 33
cggctcgagc ccagatggag gcaacagtac taggttgcag gacttgctaa gccggtaggt 60
ggcctggggt aagagatggg aagagaggct gatccctgtc ccccacagca ctggaggact 120
cccggtccca ggaaggggca aatggtgagg ccgagtcagg tgagctcagc cggcttcggg 180
ctgagctggc aggcgccctg gcagaaatgg aaaccatgaa ggctgtggca gaggtgagcg 240
agagcacgaa ggccgaggct gtggctgcgg tgcagcggca gtgccaagag gaggtggcct 300
cgctgcaggc catcctgaaa gactccatca gcagctatga agcccagatc accgccctga 360
agcaggagcg acagcagcag cagcaggact gtgaggagaa ggagcgggag ctgggccgcc 420
tgaagcagct gctgtcccgg gcctaccccc tggactccct ggagaagcag atggaaaagg 480
cccacgagga ctcggagaag ctgcgggaga tcgtactgcc catggaaaag gagatcgagg 540
agctgaaggc gaagctgctg agggccgagg agctgattca ggagatccag agacgtcccc 600
ggcatgcccc ttccctgcac ggctccacgg agttgctgcc cctgtcccgg gatccatcgc 660
ccccgctgga gcctctggag gagctgagcg gagatggggg tccagccgct gaggccttcg 720
ctcacaactg cgatgacagc gcctccatct cctccttctc ccttggcggt ggggtcggca 780
gcagctcctc cctgccccaa agccgccagg gcctgagccc tgaacaggaa gagacggcct 840
cgctggtgtc tacgggcacc ctggttcccg agggcatcta cctgccccct cctggctacc 900
agctcgtccc agacactcag tgggagcagc tgcagacaga gggccgacag ctgcagaagg 960
acctggagag cgtcagtcgc gagcgggacg agctccaaga gggcctgaga cggagcaatg 1020
aggactgtgc caagcagatg caggtgctcc tggcccaggt ccagaactca gagcagctgc 1080
tgcggaccct gcaagggacc gtgagccagg cccaggagcg ggtgcagctg cagatggcgg 1140
agctggtcac cacccacaag tgcctgcacc atgaggtaaa gcggttgaat gaggaaaacc 1200
aagggctccg ggccgagcag ctgccatcct cagcccccca gggctcgcag caggagcagg 1260
gcgaggagga atcactgccc agctctgtgc cagagctgca gcagctgctg tgctgcacgc 1320
ggcaagaggc gagggcccgg ctgcaggccc aggagcacgg ggccgagcgc ctgcggatcg 1380
agatcgtgac gctgcgggag gctctggagg aggagacagt ggccagggcc agcctggagg 1440
ggcagctgag ggtgcagcgg gaggagacag aggtgctgga ggcctccctg tgcagcctga 1500
ggacagagat ggagcgggtg cagcaggaac agagcaaggc ccagctccca gacctcctct 1560
cagaacagag ggccaaggtg ctgcggctgc aggcagagct ggagaccagt gagcaggtgc 1620
agagggattt cgtgcgactg tcccaggccc tgcaggtgcg cctagagcgg atccgccagg 1680
ctgagaccct ggagcaagtg cgcagcatca tggatgaggc gccactcacg gacgtcaggg 1740
acatcaagga cacctgaggg gtcaggatat ccccaccccc accctgggaa agacgccttt 1800
ccccactcct gaaccatgag gcctcgctct ggggtcttgg atggcttttc caccgtccct 1860
gagactgggg ttgaggggac tgacgggggc caccaccgcc ccgccctcca gcgcctcctc 1920
ccagggtggc tgggcctcct gttctcaggg atcacacctg ggtgaggggc ccaagcccct 1980
cccggaacca aaggtgcagg ctcaggcctg cggctttctg gctgctgtgc tgcctcctgg 2040
gctccagccc tcccctgccc ccagcccgtc ccctgcccag ggcacagcgg agccatgggg 2100
gctgggagtc cccatcagag gcagtgaggt gggccccggc cctgggacag gcagctgcct 2160
tctggtctgc atgacactaa gacgcttgtc cacagcggcg acccaggcct ccaagcttgc 2220
acagaggcaa ggccagactt ttccgtcgtt tattttcaat aaataagcag ctcagcgcaa 2280
aaaaaaaa 2288

<210> 34
<211> 1813
<212> DNA
<213> Homo sapiens

<220>
<221> unsure
<222> 1708, 1711, 1713, 1715
<223> Incyte ID No: a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte ID No: 3074884CB1

<400> 34
gcacgaagga ggcccggctt caggtggccc tggcggagat gccggaagat gcggacgaga 60
acgccgagga ggagctgctg cggggagagc ctctgctgcc ggcggggacc cagcgcgtgt 120
gtctggttca ccctgacgtc aagtggggcc cggggaagtc gcagatgact cgagccgagt 180
ggcaggtggc ggaggccaca gcgctggtgc acacgctgga cggctggtcc gtggtgcaga 240
caatggtcgt gtccaccaaa acgccggaca ggaagctcat ctttggcaaa gggaactttg 300
agcacctgac agaaaagatc cgagggtctc cagacgtcac gtgcgtcttc ctgaacgtgg 360
agaggatggc tgccccgacc aagaaagaac tggaagccgc ctggggcgtg gaggtgtttg 420
accgcttcac ggtcgtcctg cacatcttcc gctgtaacgc ccgcacgaag gaggcccggc 480
ttcaggtggc cctggcggag atgccgctgc acaggtcgaa cttgaaaagg gacgtcgccc 540
acctgtaccg aggagtcggc tcgcgctaca tcatggggtc aggagaatcc ttcatgcagc 600
tgcagcagcg tctcctgaga gagaaggagg ccaagatcag gaaggccttg gacaggcttc 660
gcaagaagag gcacctgctc cgccggcagc ggacgaggcg ggagttcccc gtgatctccg 720
tggtggggta caccaactgc ggaaagacca cgctgatcaa ggcactgacg ggcgatgccg 780
ccatccagcc acgggaccag ctgtttgcca cgctggacgt cacggcccac gcgggcacgc 840
tgccctcacg catgaccgtc ctgtacgtgg acaccatcgg cttcctctcc cagctgccgc 900
acggcctcat cgagtccttc tccgccaccc tggaagacgt ggcccactcg gatctcatct 960
tgcacgtgag ggacgtcagc caccccgagg cggagctcca gaaatgcagc gttctgtcca 1020
cgctgcgtgg cctgcagctg cccgccccgc tcctggactc catggtggag gttcacaaca 1080
aggtggacct cgtgcccggg tacagcccca cggaaccgaa cgtcgtgccc gtgtctgccc 1140
tgcgggggcca cggctccag gagctgaaag ctgagctcga tgcggcggtt ttgaaggcga 1200
cggggagaca gatcctcact ctccgtgtga ggctcgcagg ggcgcagctc agctggctgt 1260
ataaggaggc cacagttcag gaggtggacg tgatccctga ggacggggcg gccgacgtga 1320
gggtcatcat cagcaactca gcctacggca aattccggaa gctctttcca ggatgaacgg 1380
acgcccacag aggcctgcgg ggtgggggca tcgctgcctg gggagctgag gcgttaccgc 1440
tgtgttgggg gcagcttggt gtcaggtgca gcagggtcct ccttgtctgg ttctgcaccc 1500
gtctcgctcc cagccatttg ctgggatgac cgtgcaggcc ggtgacacgg ccgcacctgc 1560
cccaaagcgg gccgcccgag cgtccactcc aagcctgagc atccacacaa ttccagtggg 1620
ccctcggtgc ctgctgtgaa ctgctttccc tcggaatgtt tccgtaacag gacattaaac 1680
ctttgttttt acttccgtga aaaaaaanac ngngnaaaaa aaaaaggggc ggcccgctcc 1740
tagaggattc caagccttac cgtaacgcgt tgcattggcg agcggtcata agcttcttct 1800
aatagggggt cac 1813

<210> 35
<211> 1733
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3452277CB1

<400> 35
ctacggcctg gaccgagtga ccaatccgaa tgaagtcaag gtaaaccaga aacaaacagt 60
cgttgctgtc aaaaaagaga tcatgtatta ccaacaggcc ttgatgaggt ccacagtgaa 120

```
gtcttctgtg tccctgggag ggattgtgaa atacagtgag cagttctcat ccaacgatgc 180
catcatgtca ggctgcctcc ccagcaaccc ctggatcacc gatgacaccc agttctggga 240
cttaaatgcc aaattggtgg aaatcccaac caagatgcga gtggaacgat gggccttcaa 300
cttcagcgaa ttgatccgag accccaaagg tcgacagagc ttccagtact tcctcaagaa 360
agaattcagt ggagagaatc tgggattctg ggaagcctgc gaggatctga agtatggaga 420
tcagtccaaa gtcaaggaga aagcagagga gatttacaag ctgttcctgg ccccgggggc 480
gaggcgctgg atcaacatag atggcaaaac catggacatc acagtgaagg ggctgaagca 540
ccccgaccgc tatgtgctgg acgccgcaca aacccacatt tacatgctca tgaagaagga 600
ttcttatgct cgctatttaa aatctccgat ctataaggac atgctggcca aagctattga 660
acctcaggaa accaccaaga aaagctccac cctccctttt atgcggcgtc acctgcgctc 720
cagcccaagc cctgtcatcc tgagacagct ggaagaggaa gccaaggccc gagaagcagc 780
caacactgtg gacatcaccc agccgggcca gcacatggct cccagccccc atctgaccgt 840
gtacaccggg acctgcatgc cccgtctcc ttctagcccc ttctcctcct cctgccgctc 900
ccccaggaag cctttcgcct cacccagccg cttcatccgg cgacccagca ccaccatctg 960
cccctcaccc atcagagtgg ccttggagag ctcatcgggc ttggagcaga aaggggagtg 1020
cagcgggtcc atggcccccc gtgggccctc tgtcaccgag agcagcgagg cctcccctcga 1080
cacctcctgg cctcgcagcc ggcccagggc ccctcctaag gcccgcatgg ctctgtcctt 1140
cagcaggttt ctgagacgag gctgtctggc ctcacctgtc tttgccaggc tctcacccaa 1200
gtgccctgct gtgtcccacg ggagggtgca gcccctgggg gacgtgggcc agcagctgcc 1260
acgattgaaa tccaagagag tagcaaactt tttccagatc aaaatggatg tgcccacggg 1320
gagcgggacc tgcttgatgg actcggagga tgctggaaca ggagagtcgg gtgaccgggc 1380
cacagaaaag gaggtcatct gcccctggga gagcctgtaa ggaaagaggc aggctgagct 1440
gggggctctg gaccaggaag atgctctgac agatgccatg gtatgggcca caggacacac 1500
ttgctcgaga accaaagtgc atttgggtga catttgaaga ttggggagac aagatggggt 1560
agattgtggc aaagaatgct ctggctggtt accagggggcc aactccttct cctcttcctg 1620
accctccctc ccctgggcag aagaaacgca tgtggaccag aagactttcc ctgctgcctt 1680
aaaaccaata aaaggttaac tttaagtttc ttggaaaaaa aaaaaaaaag ggg        1733
```

<210> 36
<211> 1776
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4203832CB1

<400> 36

```
cccagccgag cccagcctag cccgagccca gcccgagcga accgcgagcc ccaagcccga  60
gccgcgccca gcccgagcag agccctccag ccgctcaccc cgcgtgccac cccagcgacc 120
ctcagccgct ctctgccctt ctctcggccc cgcgcccgcc ctcgcggccc ctctgcccaa 180
tgaaactggc cgcgatgatc aagaagatgt gcccgagcga ctcggagctg agtatcccgg 240
ccaagaactg ctatcgcatg gtcatcctcg gctcgtccaa ggtgggcaag acggccatcg 300
tgtcgcgctt cctcaccggc cgcttcgagg acgcctacac gcctaccatc gaggacttcc 360
accgcaagtt ctactccatc cgcggcgagg tctaccagct cgacatcctc gacacgtccg 420
gcaaccaccc gttccccgcc atgcggtgcc tctccatcct cacaggagac gttttcatcc 480
tggtgttcag tctggacaac cgcgactcct tcgaggaggt gcagcggctc aggcagcaga 540
tcctcgacac caagtcttgc ctcaagaaca aaaccaagga gaacgtggac gtgcccctgg 600
tcatctgcgg caacaagggt gaccgcgact tctaccgcga ggtggaccag cgcgagatcg 660
agcagctggt gggcgacgac ccccagcgct gcgcctactt cgagatctcg gccaagaaga 720
acagcagcct ggaccagatg ttccgcgcgc tcttcgccat ggccaagctg cccagcgaga 780
tgagcccaga cctgcaccgc aaggtctcgg tgcagtactg cgacgtgctg cacaagaagg 840
cgctgcggaa caagaagctg ctgcgggccg gcagcggcgg cggcggcggc gacccgggcg 900
acgcctttgg catcgtggca cccttcgcgc gccggcccag cgtacacagc gacctcatgt 960
```

acatccgcga gaaggccagc gccggcagcc aggccaagga caaggagcgc tgcgtcatca 1020
gctaggagcc ccgccgcgct ggcgacacaa cctaaggagg acctttttgt taagtcaaat 1080
ccaacggccc ggtgcgcccc aggccgggag cgcgcgcgga ctggcgtctc ccctccccggc 1140
gatccgcccc cagcactggg gaggcgccac tgaaccgaga agggacggtc atctgctccg 1200
gaaggaaaga gaacgggcca agactgggac tattccccac ccccggtccc ccattgaggc 1260
ccgccacccc cataactttg ggagcgaggg cccagccgag ggtggattta tcttctcaaa 1320
gacctaagag tgagcgcggg gtgggggagg gatgtgaagt tatccagcct ctgctaggct 1380
tcaagaaacc gtcatgcccg cttgagggtc aggacccacg gggcattatc ttgtctgtga 1440
ttccgggttg ctgtgacagc cggtagagcc tctgccctcc cgaaactaag cgggggggcg 1500
tgggtcaaat catagccaag tgacttgttt acatgtgagt gaaactgcac aaaggaacac 1560
aaaacaaaac ttgcacttta acggtagttc cggtgtcaac atggacacga acaaaacctt 1620
acccaggtgt ttatactgtg tgtgtgtgag gtctttaaag ttattgcttt atttggtttt 1680
ttaatataca ataaaataat ttaaaatgga aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1740
aaaaaaaaaa aaaaaaaaaa gaaaaaaaaa aaaaaa 1776

<210> 37
<211> 1316
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 104368CB1

<400> 37
cggacgcgtg ggcggcttt tctctttagc gctcagggcg ctgacccact gctcttcctc 60
ttaagaaagt gctccattcc ttccggcgcc cggagctgct ggcccaaagg gatccggagc 120
gagctagggc agacaccatg accacccttg atgataagtt gctgggggag aaactgcagt 180
actactatag cagcagtgag gatgaggaca gtgaccacga ggacaaggac cgaggcagat 240
gtgccccagc cagcagttct gtgcctgcag aggctgagct ggcaggcgaa ggcatctcag 300
ttaacacagg cccaaaaggt gtgatcaatg actggcgccg cttcaagcag ttggagacag 360
agcagaggga ggagcagtgc cgggagatgg aaaggctgat caagaagctg tcaatgactt 420
gcaggtccca tctggatgaa gaggaggagc aacagaaaca gaaagacctc caggagaaga 480
tcagtgggaa gatgactctg aaggagtttg ccataatgaa tgaggaccaa gatgatgaag 540
agtttctgca gcagtaccgg aagcagcgaa tggaagagat gcggcagcag cttcacaagg 600
ggccccaatt caagcaggtt tttgagatct ccagtggaga agggttttta gacatgattg 660
ataaagaaca gaaaagcatt gtcatcatgg ttcatattta tgaggatggc attccaggga 720
ccgaagccat gaatggttgc atgatctgcc ttgccgcaga gtacccagct gtcaagttct 780
gcaaggtgaa gagctcagtt attggcgcca gcagtcagtt caccaggaat gcccttcctg 840
ccctgctgat ctataagggg ggtgaattga tcggcaattt tgttcgtgtt actgaccagc 900
tggggggatga tttctttgct gtggaccttg aagcttttct ccaggaattt ggattactcc 960
cagaaaagga agtcttggtg ctgacatctg tgcgtaactc tgccacgtgt cacagtgagg 1020
atagcgacct ggaaatagat tgaactgata gtctagttgc atagatttct cattgtttgg 1080
gttggaatac acgtcattgt ttatttttgt tcctttgtct tctggctttt cagctgttct 1140
ttgtagtccc ttttattatg cataaaataa agaaattctt agattaaatc agaatgctga 1200
ataaccttgt agctagcaat aaggtgactt acagttgtat aaacaggaag ccaggctttt 1260
gaactgttta cttaagattc tgtggtgtga catctctgtt attgtttcca gtcaat 1316

<210> 38
<211> 1554
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1441680CB1

<400> 38
gtggcttgtg gagtggcgac cgttagtgag gcggttgctg agacagacgc tgaggcgggt 60
aggaggagcc cgagccgtaa gggaagccgt gatgagggcc gtgttgacgt ggagagataa 120
agccgagcac tgtataaatg acatcgcatt taagcctgat ggaactcaac tgattttggc 180
tgccggaagc agattactgg tttatgacac ctctgatggc accttacttc agcccctcaa 240
gggacacaaa gacactgtgt actgtgtggc atatgcgaag gatggcaagc gctttgcttc 300
tggatcagct gacaaaagcg ttattatctg gacatcaaaa ctggaaggca ttctgaagta 360
cacgcacaat gatgctatac aatgtgtctc ctacaatcct attactcatc aactggcatc 420
ttgttcctcc agtgactttg ggttgtggtc tcctgaacag aagtctgtct ccaaacacaa 480
atcaagcagc aagatcatct gctgcagctg gacaaatgat ggtcagtacc tggcgctggg 540
gatgttcaat gggatcatca gcatacggaa caaaaatggc gaggagaaag taaagatcga 600
gcggccgggg ggctccctct cgccaatatg gtccatctgc tggaaccctt caagagagga 660
acgtaatgac atcctggctg tggctgactg ggacagaaa gtttccttct accagctgag 720
tggaaaacag attggaaagg atcgggcact gaactttgac ccctgctgca tcagctactt 780
tactaaaggc gagtacattt tgctgggggg ttcagacaag caagtatctc ttttcaccaa 840
ggatggagtg cggcttggga ctgttgggga gcagaactcc tgggtgtgga cgtgtcaagc 900
gaaaccggat tccaactatg tggtggtcgg ctgccaggac ggcaccattt ccttctacca 960
gcttattttc agcacagtcc atggagttta caaggaccgc tatgcctaca gggatagcat 1020
gactgacgtc attgtgcagc acctgatcac tgagcagaaa gttcggatta aatgcaaaga 1080
gcttgtcaag aagattgcca tctacagaaa tcgattggct atccaactgc cagagaaaat 1140
cctcatctat gagttgtatt cagaggactt atcagacatg cattaccggg taaaggagaa 1200
gattatcaag aagtttgagt gcaacctcct ggtggtgtgt gccaatcaca tcatcctgtg 1260
ccaggagaaa cggctgcagt gcctgtcctt cagcggagtg aaggagcggg agtggcagat 1320
ggagtctctc attcgttaca tcaaggtgat cggtggccct cctggaagag aaggcctctt 1380
agtggggctg aagaagatgt acttgttagt gtattcattc atattgattg taaaggatta 1440
tttttcactc agtactgatg tccttggaaa tcttacctgg aaacatgttt gcaaaaaaca 1500
ttattgggtc tttcatcttt tttcttggta ttacatattt gttcaataaa aata 1554

<210> 39
<211> 2320
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1494955CB1

<400> 39
gggcgcccgc gtactcacta gctgaggtgg cagtggttcc accaacatgg agctctcgca 60
gatgtcggag ctcatggggc tgtcggtgtt gcttgggctg ctggccctga tggcgacggc 120
ggcggtagcg cgggggtggc tgcgcgcggg ggaggagagg agcggccggc ccgcctgcca 180
aaaagcaaat ggatttccac ctgacaaatc ttcgggatcc aagaagcaga aacaatatca 240
gcggattcgg aaggagaagc ctcaacaaca caacttcacc caccgcctcc tggctgcagc 300
tctgaagagc cacagcggga acatatcttg catggacttt agcagcaatg gcaaatacct 360
ggctacctgt gcagatgatc gcaccatccg catctggagc accaaggact tcctgcagcg 420
agagcaccgc agcatgagag ccaacgtgga gctggaccac gccaccctgg tgcgcttcag 480
ccctgactgc agagccttca tcgtctggct ggccaacggg gacaccctcc gtgtcttcaa 540
gatgaccaag cgggaggatg gggctacac cttcacagcc accccagagg acttccctaa 600
aaagcacaag gcgcctgtca tcgacattgg cattgctaac acagggaagt ttatcatgac 660
tgcctccagt gacaccactg tcctcatctg gagcctgaag ggtcaagtgc tgtctaccat 720

caacaccaac cagatgaaca acacacacgc tgctgtatct ccctgtggca gatttgtagc 780
ctcgtgtggc ttcaccccag atgtgaaggt ttgggaagtc tgctttggaa agaaggggga 840
gttccaggag gtggtgcgag ccttcgaact aaagggccac tccgcggctg tgcactcgtt 900
tgctttctcc aacgactcac ggaggatggc ttctgtctcc aaggatggta catggaaact 960
gtgggacaca gatgtggaat acaagaagaa gcaggacccc tacttgctga agacaggccg 1020
ctttgaagag gcggcgggtg ccgcgccgtg ccgcctggcc ctctccccca acgcccaggt 1080
cttggccttg gccagtggca gtagtattca tctctacaat acccggcggg gcgagaagga 1140
ggagtgcttt gagcgggtcc atggcgagtg tatcgccaac ttgtcctttg acatcactgg 1200
ccgctttctg gcctcctgtg ggaccgggc ggtgcggctg tttcacaaca ctcctggcca 1260
ccgagccatg gtggaggaga tgcagggcca cctgaagcgg gcctccaacg agagcacccg 1320
ccagaggctg cagcagcagc tgacccaggc ccaagagacc ctgaagagcc tgggtgccct 1380
gaagaagtga ctctgggagg gcccggcgca gaggattgag gaggagggat ctggcctcct 1440
catggcgctg ctgccatctt tcctcccagg tggaagcctt tcagaaggag tctcctggtt 1500
ttcttactgg tggccctgct tcttcccatt gaaactactc ttgtctactt aggtctctct 1560
cttcttgctg gctgtgactc ctccctgact agtggccaag gtgcttttct tcctcccagg 1620
cccagtgggt ggaatctgtc cccacctggc actgaggaga atggtagaga ggagaggaga 1680
gagagagaga atgtgatttt tggccttgtg gcagcacatc ctcacaccca aagaagtttg 1740
taaatgttcc agaacaacct agagaacacc tgagtactaa gcagcagttt tgcaaggatg 1800
ggagactggg atagcttccc atcacagaac tgtgttccat caaaaagaca ctaagggatt 1860
tccttctggg cctcagttct atttgtaaga tggagaataa tcctctctgt gaactccttg 1920
caaagatgat atgaggctaa gagaatatca agtccccagg tctggaagaa aagtagaaaa 1980
gagtagtact attgtccaat gtcatgaaag tggtaaaagt gggaaccagt gtgctttgaa 2040
accaaattag aaacacattc cttgggaagg caaagttttc tgggacttga tcatacattt 2100
tatatggttg ggacttctct cttcgggaga tgatatcttg tttaaggaga cctcttttca 2160
gttcatcaag ttcatcagat atttgagtgc ccactctgtg cccaaataaa tatgagctgg 2220
ggattaaata cgaataagac atggtttctg ccatcaaaga tggctggtgg gagagagaga 2280
tacaccctta ttaagtgctt tgtgttagtt tattcatagc 2320

<210> 40
<211> 879
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1508161CB1

<400> 40
gaagaatttg ttcaggcgtt cgtgcagaag gacccctttgg ataatgacaa gagttgctac 60
agtgaacgga agaaaaacacg aaacttagaa gcttacgtgg aatggtttaa tcgcctcagc 120
tacttggttg ctacagaaat ctgtatgcct gttaagaaaa aacaccgagc aagaatgatt 180
gagtatttca ttgacgtagc tcgggagtgt tttaacattg gcaacttcaa ctccttgatg 240
gcgataatct ctggtatgaa tatgagccca gtctctcgac taaaaaaaac ttgggccaaa 300
gtgaagactg caaaatttga cattcttgag catcagatgg acccttcaag caatttctat 360
aattatcgaa cagctcttcg tggggcagca caaaggtctt taactgctca tagtagtaga 420
gaaaagattg tgataccatt cttcagtctc ttaatcaaag atatttattt cctcaatgag 480
ggttgtgcca accgccttcc caatggccat gtcaattttg agaaattttg ggaactggcc 540
aaacaagtga gtgaatttat gacatggaaa caagtggagt gtccatttga gagggaccgg 600
aagatcttgc agtatctgct cacagtacca gtcttcagtg aagatgctct ctacttggct 660
tcttatgaga gtgaaggacc tgaaaatcat atagagaaag acagatggaa gtctttaagg 720
tcgagcctct taggcagagt ttaacacatg ggagctgcct gcctgctgct gctgctgctt 780
cctgcagatc atggaggggc tggcctttgt tttctggcat ctcgtaccac gaacactcat 840
gaagaccctg cagtcattgg agcacccggg tcagcaaag 879

<210> 41
<211> 2248
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1811877CB1

<400> 41
ggaggaccag gaggacatca ctgcctttga cctcagccct gacaacgagg tgctggtgac 60
agccagtcgg gcattgctgc tggctcagtg ggcctggcaa gagggcagcg ttacccgcct 120
gtggaaggcg atacacacgg cccccgtggc caccatggcc ttcgacccca cctccactct 180
gctagccaca ggtggctgtg atggggccgt gcgcgtctgg gacatcgtgc ggcactacgg 240
gacacaccac ttccgaggct cgccccggtgt cgtgcaccta gtggccttcc acccggaccc 300
tacacgcctg ctgctcttct cctcggccac ggatgccgcc atccgcgtgt ggtcactgca 360
ggaccggtca tgcctggctg tgctgactgc ccactacagc gccgtcacct cactggcctt 420
cagcgccgac ggccacacca tgctcagctc cggccgtgac aagatatgta tcatctggga 480
ccttcagagc tgccaggcca cgaggaccgt gcctgtgttt gagagcgtgg aggctgctgt 540
gctgttgcca gaggagccag tgtcccagct gggtgtgaag tccccagggc tgtactttct 600
gacagctggc gaccaaggca ctctgcgcgt gtgggaggca gcttctgggc agtgtgtgta 660
cacgcaggcc cagccgccgg gccctgggca ggagctgacc cactgcaccc tggcacacac 720
cgccggcgtg gtcctcaccg ccaccgccga ccacaacctg ttgctctacg aggctcgctc 780
cctgcggctg cagaaacagt tcgctggcta cagtgaggag gttttggatg tccggtttct 840
tgggcccgag gactcccacg ttgtcgtggc ctccaatagc ccctgcctaa aagtgtttga 900
gctgcagacg tcagcctgcc agatcctcca cggccacacg gatatcgtcc tggccctgga 960
tgtgttccgg aaggggtggc tctttgccag ctgtgccaag gatcagagcg tccgtatctg 1020
gagaatgaac aaggctggcc aggtgatgtg cgtggctcag ggttccggtc acacacacag 1080
tgtgggcacc gtctgctgct ctaggctgaa ggagtccttc ctggtgacag gcagccagga 1140
ctgcactgtg aagctgtggc ctcttcccaa agccttgctg tccaagaaca cagccccaga 1200
caacggccct atcctcctgc aggcccagac cactcagcgc tgccatgata aggacatcaa 1260
cagcgtggct attgcccccca acgacaagct gctggccaca ggctcacagg accgcacggc 1320
caagctctgg gccctgccac agtgccagct gctgggtgtc ttctcaggcc accggcgtgg 1380
cctctggtgc gtccagttct ctcccatgga ccaggtgctg gccacggcct cagctgatgg 1440
caccatcaag ctctgggcac tccaggactt cagctgtctc aagacatttg aggggcacga 1500
tgcttctgtg ctgaaggtgg cctttgtgag ccgtggcacg cagctgctgt ccagcggttc 1560
ggatggcctc gtgaagctct ggaccatcaa gaacaacgag tgtgtgcgga cgctggatgc 1620
ccacgaggac aaggtctggg ggctgcactg cagccggctg gacgaccacg ccctcactgg 1680
ggccagtgac tcccgagtca tcctctggaa ggatgtgacc gaggcggagc aggcagagga 1740
gcaggccagg caagaggagc aggtggtcag gcagcaagag ctggacaacc tgctgcatga 1800
gaagcggtac ctgcgggcgc tgggcctggc catctccctg gatcggcccc acaccgtgct 1860
gactgtcatc caggccatcc ggagggaccc tgaggcctgc gagaagctgg aagccaccat 1920
gctccgactg cggcgcgacc agaaagaggc cctgctgcgc ttctgcgtca cgtggaacac 1980
caactcgcgg cactgccacg aggcccaggc cgtgctgggt gtgctcttga ggcgagaggc 2040
ccccgaggag ctgctggcct acgaaggcgt gcgggcagcg cttgaggccc tgctgcccta 2100
cactgagcgg cactttcagc ggctcagcag gaccctccag gccgccgctt tcttggactt 2160
cctgtggcac aacatgaagc tccctgtgcc ggccgccgcc cccacccccct gggaaaccca 2220
taaaggcgca ctgccctaaa aaaaaaaa 2248

<210> 42
<211> 2146
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1848674CB1

<400> 42
gttattggca agttcccctg cagttgtttg ggctgtccct gtggctggtt ctggggtgtg 60
cggccagcca tggagcgctc tgggcccagc gaagtgacag gctcagacgc gtcgggaccg 120
gacccgcagc ttgcggtcac catgggcttc acggggttcg gtaaaaaagc tcgcacattt 180
gacttggaag caatgtttga acaaactcga aggacagctg tggaaagaag tcgcaaaaca 240
ctggaagcaa gagaaaaaga ggaagaaatg aacagagaga aagaattaag aagacaaaat 300
gaagatattg agccaacatc ctcaagatca aatgtggtca gagattgctc caaatcatct 360
tccagggata cgagcagcag tgaaagtgaa cagagttctg actcttctga tgatgagtta 420
attggccctc ctttaccccc taaaatggta ggaaaaccag ttaattttat ggaggaagat 480
atcctcggtc ctttacctcc acctcttaat gaagaagaag aagaagcaga ggaagaagaa 540
gaggaagagg aggaagagga aaatcctgtt cacaagattc ctgactcgca tgagataacg 600
ctgaagcatg gcactaaaac agtgtctgct ttgggtctgg atccctcagg tgcccgtttg 660
gtgacaggag gatatgacta tgatgttaag ttttgggatt ttgctggaat ggatgcttct 720
tttaaggcat ttcgatccct tcagccctgt gagtgccatc agatcaagtc attacagtat 780
agtaacacag gagacatgat tcttgttgta tctggaagct ctcaggccaa ggtgattgac 840
agagatggtt ttgaagtaat ggaatgtata aaaggagacc agtatattgt ggacatggcc 900
aacaccaagg gtcatacagc aatgcttcat actggctcat ggcatcccaa aataaaggga 960
gaatttatga cttgctcaaa tgatgcgact gtgaggacgt gggaagttga aaatccaaag 1020
aagcaaaaaa gtgtgtttaa accacggacg atgcaaggca aaaaagtcat tcccactacg 1080
tgcacatata gtagagatgg aaacctcata gcagctgcct gccagaatgg aagcatacag 1140
atctgggacc gaaatttgac tgttcatcct aagttccact ataaacaggc tcatgactcg 1200
ggcacagaca cttcttgcgt gactttttcc tatgatggta atgtccttgc ctctcgtgga 1260
ggtgacgatt cattaaaatt atgggacatc cgacaattta ataaaccact tttttcagcc 1320
tcgggtcttc ccaccatgtt cccaatgact gactgctgtt tcagtccaga tgataagctc 1380
atagtcactg gtacatctat tcaaagagga tgtggcagcg gcaaacttgt tttctttgag 1440
cgtaggactt tccaaagggt gtatgaaata gacatcacag atgcgagtgt tgttcgctgc 1500
ctgtggcatc caaagctgaa ccagatcatg gttggaactg gaaatggatt ggctaaagtc 1560
tattacgacc ccaacaagag tcagaggggа gcaaaattat gtgtggttaa aacccagcgg 1620
aaggcaaaac aagctgagac tctaactcag gactacatca tcacccctca tgccttgcct 1680
atgttccgtg agccccgcca acggagtaca aggaaacagc tggagaagga cagactggat 1740
cccctgaagt cgcataaacc tgaacctcct gtagcaggcc caggtcgtgg tggccgagtt 1800
ggaacccacg gggcactctc ctcttcctat attgtgaaga acattgcttt ggacaagacc 1860
gatgacagta atcctcggga agccattttg cgtcatgcca aagcagcaga agacagccca 1920
tattgggttt ctccagcata ttccaagact cagcccaaaa ccatgtttgc ccaagttgaa 1980
tctgatgatg aggaagcaaa gaatgagcca gaatggaaaa aacgtaaaat ttgaagaatc 2040
tcatttgaga gctgtttgca tgagtgggag gggtatggga caggtttggg tttttttttt 2100
atgctcatga aattaaaaat tcatttttat gaaaaaaaaa aaaaaa 2146

<210> 43
<211> 714
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2012970CB1

<400> 43
ggatggcgag cagcggaggc gaggcggtga cgagagcagc ggctccgcca ttggacgagg 60

```
aggcctgagg gacgggccag cggtgcacaa gaagagaccg aggcgggtgg ccccgagaga 120
gccagggcca tggaggccaa catgccgaag cggaaggagc ctggcaggtc tctccgtatc 180
aaagtcatct ccatgggcaa cgccgaagtg gggaaaagct gtattataaa gcgatactgt 240
gagaaaagat tcgtgtctaa atacctggca acaattggaa ttgactatgg agtcacaaag 300
gtacacgtca gagacagaga aatcaaagtt aacatctttg atatggctgg acatcccttc 360
ttctatgagg ttcgaaatga gttttacaag gacacacagg gtgtgatact ggtctatgat 420
gttgggcaga aagactcctt tgacgccctt gatgcgtggc tggcagaaat gaagcaagag 480
cttggacctc atggaaacat ggaaaatatt atatttgtag tttgtgccaa caagattgat 540
tgtaccaaac atcgctgtgt agatgaaagt gaaggacgtc tttgggctga aagcaaaggg 600
ttcctgtact ttgaaacttc agcacaaact ggagaaggca ttaatgagat gttccagata 660
catcttggat agaactaatg gataaattag tctgtttaaa aaagaaaaaa aaaa       714
```

<210> 44
<211> 1779
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2254315CB1

<400> 44

```
caggaggaag atggcggcgt ccgcagctgc cgctgagctc caggcttctg gggtccgcg   60
gcacccagtg tgtctgttgg tgttgggaat ggcgggatcc gggaaaacca cttttgtaca 120
gaggctcaca ggacacctgc atgcccaagg cactccaccg tatgtgatca acctggatcc 180
agcagtacat gaagttccct ttcctgccaa tattgatatt cgtgatactg taaagtataa 240
agaagtaatg aaacaatatg gacttggacc caatggcggc atagtgacct cactcaatct 300
ctttgctacc agatttgatc aggtgatgaa atttattgag aaggcccaga acatgtccaa 360
atatgtgttg attgacacac ctggacagat tgaggtattc acctggtcag cttctgggac 420
aattatcact gaagcccttg catcctcatt tccaacagtt gtcatctatg taatggacac 480
atcgagaagt accaacccag tgaccttcat gtccaacatg ctctatgcct gcagcatctt 540
atacaaaacc aagctgcctt tcattgtggt catgaataaa actgacatca ttgaccacag 600
ctttgcagtg gaatggatgc aggattttga ggctttccaa gatgccttga atcaagagac 660
tacatacgtc agtaacctga ctcgttcaat gagcctggtg ttagatgagt tttacagctc 720
actcagggtg gtgggtgtct ctgctgttct gggtactgga ttagatgaac tctttgtgca 780
agttaccagt gctgccgaag aatatgaaag ggagtatcgt cctgaatatg aacgtctgaa 840
aaaatcactg gccaacgcag agagccaaca gcagagagaa caactggaac gccttcgaaa 900
agatatgggt tctgtagcct tggatgcagg gactgccaaa gacagcttat ctcctgtgct 960
gcacccttct gatttgatcc tgactcgagg aaccttggat gaagaggatg aggaagcaga 1020
cagcgatact gatgacattg accacagagt tacagaggaa agccatgaag agccagcatt 1080
ccagaatttt atgcaagaat cgatggcaca atactggaag agaaacaata aataggagac 1140
tttagcacac ttcacttgtt tctagaagtc cagaatttg gacctccacg tgaaagaact 1200
gttcttacct ctgaactggg ggctcccata agggataatt ttcctcagag tagcaaagtt 1260
tctcttatta gagaaatctt gtgactcaga tgaagtcagg gatagaagac ccttggacct 1320
ggcaggttaa tgctgattat tccttggcct ttcccttgta tttatgcaag gaaggatata 1380
ctgagctgat actcttccaa gcctacaact tcaagtttta tcatttgaac tcaagtactt 1440
ttgctgctga ggaatggaat caaaagaacg tagtctcctg gtgaccacct cagatctcta 1500
ttattaggct agatgtatag cctctactcc cccagcttct tgctcttgac cctgcactgt 1560
aagttgccct tctattagca gccaaggaaa agggaaacat gagcttatcc agaacggtgg 1620
cagagtctcc ttggcaatca accaacgttg ctatgaaata tgcctcacac tgtatagctc 1680
attataggac gtcaggtttg ttgaaaaaag tgggcaagac atgattaatg aatcagaatc 1740
ctgtttcatt ggtgacttgg ataaagactt tttaatttt                        1779
```

<210> 45
<211> 2234
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2415545CB1

<400> 45
```
ccccttctga aaaatggttt cacccgctcc ttctctagaa cgggtcgggc cgctaccact   60
gtccggcccg gaggggaact gttttctccg gaagtgacaa cacgctgact aggaaaagga  120
ggaggcgggg cagtggggcc ttcggcggcg actatggaag gagccggcta caggtggtg  180
tttgagaagg gcggagtgta cctgcacacc agcgctaaga agtatcagga ccgagactct  240
ctcatcgctg gtgtcatccg tgtcgtggaa aaggacaatg acgtcctcct gcactgggct  300
cctgtagagg aggctggaga ttccacccaa atcctcttct ccaagaagga ctccagtggg  360
ggtgactcat gtgcttctga ggaggaacca acctttgacc ccggctatga acctgactgg  420
gctgtcatca gcactgtgcg gccacagccc tgccactcag agcccacgag aggtgcagag  480
cccagctgcc cccagggctc ctgggccttc tcagtgagtc tgggggagct aaagtccatc  540
cgccgctcca agccaggcct cagctgggcc tacctggttc tggtgaccca ggctggaggt  600
tccctgcccg cactgcactt ccaccgcggg ggcacccgcg ccctgctccg cgtcctcagc  660
cgctacctgc tgttggccag ctccccgcag gactcccgcc tctaccttgt cttcccccac  720
gactcctctg ctctctccaa ctccttccac cacctgcagc tctttgacca ggacagctcc  780
aatgtggtgt cacgcttcct ccaggatccc tactccacca ccttcagcag cttctcccga  840
gtgaccaact tcttccgggg tgccctgcag ccacagcctg agggagccgc ctccgacctt  900
cccccgccac ccgacgatga gcccgagcct ggattcgagg tcatttcctg tgtggagctg  960
gggcctcggc caaccgtgga gcggggccct ccagttacag aggaggagtg ggcacgccac 1020
gtgggccctg aaggtcgcct gcagcaggtc cctgagctga agaaccggat cttctcgggg 1080
ggtctgagcc ccagcctgcg gcgcgaggcc tggaagttcc tcctagggta cctcagctgg 1140
gaaggcacag ctgaggagca caaggcccac atacgcaaga aaacggatga gtatttccgc 1200
atgaagctgc agtggaaatc tgtgagccct gagcaggagc ggagaaactc acttctgcat 1260
ggataccgca gcctcatcga aagggatgtg agccgcactg acaggaccaa caagttctac 1320
gagggtcccg agaacccggg gctgggcctg ctgaacgata tcctcctcac ctactgcatg 1380
tatcacttcg acctcggcta cgtccagggc atgagtgatc ttctctcccc gatcctctac 1440
gtcattcaga acgaggtgga tgctttctgg tgtttctgtg gcttcatgga gctcgtgcaa 1500
gggaactttg aagagagcca ggagaccatg aagcggcaac tcgggcgact gctgctgctc 1560
ctgagggtgc tggacccccct gctctgcgac ttcctggatt cccaggactc cggctctctc 1620
tgcttctgtt tccggtggct gctcatctgg ttcaagaggg aattcccctt cccggatgtc 1680
cttcggctgt gggaggtgct gtggacaggg ctccctggcc ccaatctgca cctgctggtg 1740
gcctgcgcca tcctggacat ggagagggac accctcatgc tgtccggctt cggctccaat 1800
gagatcctca agcacatcaa cgagctgact atgaagctga gcgtggagga cgtgctgacc 1860
cgcgccgagg ccctgcaccg ccagctaacc gcctgcaccc gagctgcccc acaacgtgca 1920
ggagatcctg ggctggcccc cgccacgcag agccccacag cccctcgccc accgcctccc 1980
cgctgcctct gtacgcccac ccgggccccg cccaccccgc cgccctccac ggacacagcc 2040
ccgcagcccg acagcagcct ggagatcctg cccgaggagg aggacgaggg cgccgactcc 2100
taaccccgcc aggcagcctc gttctgcaca ggcactttag cccgagccag gcacacctgc 2160
gagggggcag gtgtgctccg ccgccctgct gataagctgg cttcattaaa ctgacacttc 2220
tcaaaaaaaa aaaa                                                   2234
```

<210> 46
<211> 3150
<212> DNA
<213> Homo sapiens

<220>
<221> unsure
<222> 96, 97, 99, 3070-3072, 3074, 3078, 3080-3082, 3085-3087, 3091,
<222> 3099, 3100, 3103, 3107, 3110-3112, 3114, 3115, 3121, 3123, 3125,
<222> 3128, 3136, 3138, 3140, 3141, 3143, 3145, 3147, 3149
<223> Incyte ID No: a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte ID No: 2707969CB1

<400> 46
acacaggagc aatgcaaaat ttgataaagc atttttctgt cagatcagct gagccctact 60
gcccttcctc tcaagattcc tggagaccca gatgtnngna tctttcattg acaacaaaat 120
aatgtgtcat gatgatgatg ataaagaccc tgtactccgg gtatttgatt cccgagttga 180
caagatcagg ctgttgaatg ttcggacacc tactctccgt acatccatgt accagaagtg 240
taccactgtg gatgaagcag agaaagcaat tgagctgcgt ctggcaaaaa ttgaccatac 300
tgcaattcac ccacatttac ttgacatgaa gattggacaa gggaaatatg agccgggctt 360
cttccctaag ctgcagtctg atgtactttc cactgggcca gccagcaaca agtggacgaa 420
aaggaatgcc cctgcccagt ggaggcggaa agatcggcag aagcagcaca cagaacacct 480
gcgtttagat aatgaccaga gggagaagta catccaggaa gccaggacta tgggcagcac 540
tatccgccag cccaaactgt ccaacctctc tccatcagtg attgcccaga ccaattggaa 600
gtttgtagag ggcctgctga aggaatgccg caataagacc aagaggatgc tggtggaaaa 660
gatgggccga gaagctgtgg agctagggca tggggaggtg aacatcacag ggtggaaga 720
gaacaccctg attgccagcc tttgtgatct cctggaaagg atctggagtc atggactaca 780
agtgaaacag gggaaatcag ccttatggtc ccacctgtta cattatcagg acaaccggca 840
gagaaaactc acatcaggaa gcctcagtac ctcaggaata cttcttgatt cagaacgtag 900
gaagtctgat gccagctcac tcatgcctcc cctgaggatc tccctgattc aggatatgag 960
gcacatccag aacatcgggg aaatcaagac tgatgtggga aaggccagag catgggtgcg 1020
actgtccatg gaaaaaaagt tactttccag acacctgaag cagctcctct cagaccatga 1080
gctcaccaaa aagttatata agcgctatgc cttcctgcgc tgtgatgacg agaaggagca 1140
gttcctctat cacctcctgt ctttcaatgc cgtcgattac ttttgcttca ccaatgtctt 1200
cacaactatc ctgatcccgt accacattct gatcgtacca agcaagaagc tgggggggctc 1260
catgttcact gccaacccat ggatctgtat atcaggagaa ttgggtgaga cacagatcat 1320
gcagattccc aggaatgtgc tagagatgac cttcgagtgc cagaacttgg ggaagcttac 1380
tactgtccag attggccatg ataactctgg gctgtatgcc aaatggctgg tggagtatgt 1440
gatggtcagg aatgagatca caggacatac ctacaagttc ccgtgtggcc ggtggttagg 1500
gaagggcatg gatgatggaa gcctggagcg gatcctagtt ggggagctgc tcacatccca 1560
gcctgaggtg gatgagaggc catgccggac cccgccgctg cagcagtccc ccagtgtcat 1620
ccggaggctt gttaccatct cacccaacaa caagcccaag ctgaacactg ggcagatcca 1680
ggagtccatc ggggaggcag tcaatggcat tgtgaagcac ttccataagc ctgagaaaga 1740
gcgaggcagt ctgacgctgt tgctctgtgg agagtgtggc cttgtctcgg ccttggaaca 1800
ggctttccag catggattta aatcgccccg gctcttcaaa aatgtcttca tttgggattt 1860
cctggaaaaa gcacaaacct attatgagac attagagaag aatgaagtag tccctgagga 1920
aaactggcat acaagagccc ggaacttctg ccgatttgtc actgcaatca acaatactcc 1980
ccggaacatc ggcaaggatg gcaagtttca gatgctggtg tgcttgggag ccagagatca 2040
cctcctacac cactggattg ccctgctggc tgactgcccc atcactgcac acatgtatga 2100
ggatgtggca ctgatcaaag accatacact tgtcaattcc ttgattcgtg tgctgcagac 2160
attgcaggag ttcaacatca cgctggagac gtcccttgtc aagggcatcg acatctgacc 2220
tcccagcacc agccagcagc aggactgaga aagactcacc ctgcagctct gaccttttttt 2280
cccaaaggga cttaagcgat tgtgcaggag taggagacaa aatgtacact cactgtaaaa 2340
agagaactag aggatttttg gaataaataa tctattttag agtttatttg ctgatttgct 2400
ttttacacac tttcatgtga aagagtgata gggagaggga gcgaggctgg tgccgcttat 2460
tttgaagctg gtgccctccc tcgccgtggc cacatgctgg aagcctgagg cctccctgga 2520
ctgagcctgt ggcactgcgt gcgggacagt tatgtttcct tgccccgtcg cattaatgag 2580

gcccttccac atcattttta aactaatgtt tttctatatt aacattatta tggatatttg 2640
gctttcatag gccacacaca ggtgtgctgc gcgggaagcc ccatgctcca atcaaaggga 2700
tttttagtag tgcctctaag caagcaccga tgagtcagtc ccacgtattt tctttttttgt 2760
cagtattgtt tgggaaggag acatgccggg atgtgtcatc gtgccaaata ccacatttcc 2820
tgttggcaca gtttcacaga agtaaacata agcatgtttt aacaggtttt tcttttcttt 2880
tttctttttt aaaatgtttt atttatttaa cccgccattg tgtgttttta agtattttct 2940
ttttttaagg aaaggaaaag cttgtcacaa tctaactggc tatgttatta ttattaaatt 3000
tatgttttgc aacttagaaa ccagctacag tatggcccac ttaataaaac acctgaaaca 3060
aaaaaaaagn nngngggngn nngtnnngag naggagggnn ggngggnggn nngnngggag 3120
ntnanttntg gggtgngngn ngngnangnt 3150

<210> 47
<211> 1806
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2817769CB1

<400> 47
gtctgcgcgc aggtgccgct cggcgcccgg gccgcccgtt ccgcggctgt cgccgccgtc 60
gtgcgtgccg ctcggcggag gggacgggcc tgcgttctct cctccttcct ccccgcctcc 120
agctgccggc aggacctttc tctcgctgcc gctgggaccc cgtgtcatcg cccaggccga 180
gcacgatgcc ccctaaaaag ggaggtgatg gaattaaacc acccccaatc attggaagat 240
ttggaacctc actgaaaatt ggtattgttg gattgccaaa tgttgggaaa tctactttct 300
tcaatgtgtt aaccaatagt caggcttcag cagaaaactt cccgttctgc actattgatc 360
ctaatgagag cagagtacct gtgccagatg aaaggtttga ctttctttgt caataccaca 420
aaccagcaag caaaattcct gcctttctaa atgtggtgga tattgctggc cttgtgaaag 480
gagctcacaa tgggcagggc ctggggaatg ctttttttatc tcatattagt gcctgtgatg 540
gcatctttca tctaacacgt gcttttgaag atgatgatat cacgcacgtt gaaggaagtg 600
tagatcctat tcgagatata gaaataatac atgaagagct tcagcttaaa gatgaggaaa 660
tgattgggcc cattatagat aaactagaaa aggtggctgt gagaggagga gataaaaaac 720
taaaacctga atatgatata atgtgcaaag taaaatcctg ggttatagat caaaagaaac 780
ctgttcgctt ctatcatgat tggaatgaca aagagattga agtgttgaat aaacacttat 840
ttttgacttc aaaaccaatg gtctacttgg ttaatctttc tgaaaaagac tacattagaa 900
agaaaaacaa atggttgata aaaattaaag agtgggtgga caagtatgac ccaggtgctt 960
tggtcattcc ttttagtggg gccttggaac tcaagttgca agaattgagt gctgaggaga 1020
gacagaagta tctggaagcg aacatgacac aaagtgcttt gccaaagatc attaaggctg 1080
ggtttgcagc actccaacta gaatactttt tcactgcagg cccagatgaa gtgcgtgcat 1140
ggaccatcag gaaagggact aaggctcctc aggctgcagg aaagattcac acagattttg 1200
aaaagggatt cattatggct gaagtaatga aatacgaaga ttttaaagag gaaggttctg 1260
aaaatgcagt caaggctgct ggaaagtaca gacaacaagg cagaaattat attgttgaag 1320
atggagatat tatcttcttc aaatttaaca cacctcaaca accgaagaag aaataaaatt 1380
tagttattgc tcagataaac atacaacttc caaaaggcat ctgattttta aaaaattaaa 1440
atttctgaaa accaatgcga caaataaagt tggggagatg ggaatctttg acaaacaaat 1500
tatttttatt tgttttaaaa ttaaaatact gtgtaccccc ccccactcca tgaaatgcag 1560
gttcactaaa tgtgaacagc tttgcttttc acgtgattaa gaccctactc caaattgtag 1620
aagcttttca ggaaccatat tactctcatg atacttcatt aatctccatc atgtatgcca 1680
agcctgacac atttgacagt gaggacaatg tggcttgctc ctttttgaat ctacagataa 1740
tgcatgtttt acagtactcc agatgtctac actcaataaa acatttgaca aaaccaaaaa 1800
aaaaaa 1806

<210> 48
<211> 2880
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2917557CB1

<400> 48
```
gaggaggagg aggaggaaga agaggaagaa gaagatgaag aaagtgaaga agaggaggaa   60
gaggagggag aaagtgaagg cagtgaaggt gatgaggaag atgaaaaggt gtcagatgag  120
aaggattcag ggaagacatt agataaaaag ccaagtaaag aaatgagctc agattctgaa  180
tatgactctg atgatgatcg gactaaagaa gaaagggctt atgacaaagc aaaacggagg  240
attgagaaac ggcgacttga acatagtaaa aatgtaaaca ccgaaaagct aagagcccct  300
attatctgcg tacttgggca tgtggacaca gggaagacaa aaattctaga taagctccgt  360
cacacacatg tacaagatgg tgaagcaggt ggtatcacac aacaaattgg ggccaccaat  420
gttcctcttg aagctattaa tgaacagact aagatgatta aaaattttga tagagagaat  480
gtacggattc caggaatgct aattattgat actcctgggc atgaatcttt cagtaatctg  540
agaaatagag gaagctctct ttgtgacatt gccattttag ttgttgatat tatgcatggt  600
ttggagcccc agacaattga gtctatcaac cttctcaaat ctaaaaaatg tcccttcatt  660
gttgcactca ataagattga taggttatat gattggaaaa agagtcctga ctctgatgtg  720
gctgctactt taaagaagca gaaaaagaat acaaaagatg aatttgagga gcgagcaaag  780
gctattattg tagaatttgc acagcagggt ttgaatgctg ctttgtttta tgagaataaa  840
gatccccgca cttttgtgtc tttggtacct acctctgcac atactggtga tggcatggga  900
agtctgatct accttcttgt agagttaact cagaccatgt tgagcaagag acttgcacac  960
tgtgaagagc tgagagcaca ggtgatggag gttaaagctc tcccggggat gggcaccact 1020
atagatgtca tcttgatcaa tgggcgtttg aaggaaggag atacaatcat tgttcctgga 1080
gtagaagggc ccattgtaac tcagattcga ggcctcctgt tacctcctcc tatgaaggaa 1140
ttacgagtga agaaccagta tgaaaagcat aaagaagtag aagcagctca gggggtaaag 1200
attcttggaa aagacctgga gaaaacattg gctggtttac ccctccttgt ggcttataaa 1260
gaagatgaaa tccctgttct taaagatgaa ttgatccatg agttaaagca gacactaaat 1320
gctatcaaat tagaagaaaa aggagtctat gtccaggcat ctacactggg ttctttggaa 1380
gctctactgg aatttctgaa aacatcagaa gtgccctatg caggaattaa cattggccca 1440
gtgcataaaa aagatgttat gaaggcttca gtgatgttgg aacatgaccc tcagtatgca 1500
gtaattttgg ccttcgatgt gagaattgaa cgagatgcac aagaaatggc tgatagttta 1560
ggagttagaa tttttagtgc agaaattatt tatcatttat ttgatgcctt tacaaaatat 1620
agacaagact acaagaaaca gaaacaagaa gaatttaagc acatagcagt atttccctgc 1680
aagataaaaa tcctccctca gtacattttt aattctcgag atccgatagt gatgggggtg 1740
acggtggaag caggtcaggt gaaacagggg acacccatgt gtgtcccaag caaaaatttt 1800
gttgacatcg gaatagtaac aagtattgaa ataaaccata aacaagtgga tgttgcaaaa 1860
aaaggacaag aagtttgtgt aaaaatagaa cctatccctg gtgagtcacc caaaatgttt 1920
ggaagacatt ttgaagctac agatattctt gttagtaaga tcagccggca gtccattgat 1980
gcactcaaag actggttcag agatgaaatg cagaagagtg actggcagct tattgtggag 2040
ctgaagaaag tatttgaaat catctaattt tttcacatgg agcaggaact ggagtaaatg 2100
caatactgtg ttgtaatatc ccaacaaaaa tcagacaaaa aatggaacag acgtatttgg 2160
acactgatgg acttaagtat ggaaggaaga aaaataggtg tataaaatgt tttccatgag 2220
aaaccaagaa acttacactg gtttgacagt ggtcagttac atgtccccac agttccaatg 2280
tgcctgttca ctcacctctc ccttccccaa cccttctcta cttggctgct gttttaaagt 2340
ttgcccttcc ccaaatttgg atttttatta cagatctaaa gctctttcga ttttatactg 2400
attaaatcag tactgcagta tttgattaac caagcttctg cagattttgt gattcttggg 2460
actttttga cgtaagaaat acttctttat ttatgcatat tcttcccaca gtgatttttc 2520
cagcattctt ctgccatatg cctttagggc ttttataaaa tagaaaatta ggcattctga 2580
tatttcttta gctgctttgt gtgaaaccat ggtgtaaaag cacagctggc tgctttttac 2640
tgcttgtgta gtcacgagtc cattgtaatc atcacaattc taaaccaaac taccaataaa 2700
```

gaaaacagac atccaccagt aagcaagctc tgttaggctt ccatgttagt gtagcttctc 2760
tcccacaagt tgtcctccta ggacaagaat tatcttacaa actaaactat catcacacta 2820
ccttgtatgc cagcacctgg taacagtaga gatttttata cattaatctt gatctgtttt 2880

<210> 49
<211> 1109
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3421335CB1

<400> 49
cccacgcgtc cgctcgctct tgggtcatgc ctggccagca gaaagcagct ccatagggga 60
ggagagccac gcaggatctc acagctgcag tctaatagta acacagagga ttcagcagtg 120
gccaccatgg gttctgtgaa ttccagaggt cacaaggcgg aagcccaggt ggtgatgatg 180
ggcctggact cggcgggcaa gaccacgctc ctttacaagc tgaagggcca ccagctggtg 240
gagaccctgc ccactgttgg tttcaacgtg gagcctctga aagctcctgg gcacgtgtca 300
ctgactctct gggacgttgg ggggcaggcc ccgctcagag ccagctggaa ggactatctg 360
gaaggcacag atatcctcgt gtacgtgctg gacagcacag atgaagcccg cttacccgag 420
tcggcggctg agctcacaga agtcctgaac gaccccaaca tggctggcgt ccccttcttg 480
gtgctggcca acaagcagga ggcacctgat gcacttccgc tgcttaagat cagaaacagg 540
ctgagtctag agagattcca ggaccactgc tgggagctcc ggggctgcag tgccctcact 600
ggggaggggc tgcccgaggc cctgcagagc ctgtggagcc tcctgaaatc tcgcagctgc 660
atgtgtctgc aggcgagagc ccatggggct gagcgcggag acagcaagag atcttgatcc 720
agacagagca gcatatcttt gctcatacaa actagaagaa ccagctgatc cttgagaaat 780
ttacgcttag tctatcaaac aagaaatgct ggcttggccc ggtggctcat gcctgtaatc 840
ccagcactgt gggagaccac ggtgggggaa tcccttgagc ccaggagttg gagagcaaca 900
tcacaacacc ccatttctac taataatcaa aaaattggcc gggcatggtg gcatgtgcct 960
gtagtcccag ctacttggga ggctgaggca ggagaatcgc ttgagcccaa gaggtagagg 1020
ttgcagtgag ccaagatcgc gccactgcac tccagtctgg gcaacagagt gagaccctgt 1080
tctagtggtg ataataataa tgatgtagt 1109

<210> 50
<211> 2407
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 605761CB1

<400> 50
ctttcgctgg cgccattacc tgagttctcc tccagcgttt ccgcaccctc tccgattagc 60
ggtcccagga gtttccaagg taaccgcgca gtagggcgga tctcattagg cggaaagcga 120
aacccggaag tgacgctctt accgggtgtc agcagcgaga gggttcgaag atggcggcgc 180
gcaagggtcg gcgtcgcacg tgtgaaaccg gggaacccat ggaagccgag tccggcgaca 240
caagttccga gggcccggcc caggtctacc tgcccggccg ggggccgccg ctacgcgaag 300
gggaggagct ggtcatggac gaggaggcct atgtgctcta ccaccgagcg cagactggcg 360
ccccctgtct cagctttgac atagtccggg atcacctggg agacaaccgg acagagcttc 420
ctcttacact ttacttgtgt gctgggaccc aggctgagag cgcccagagc aacagactga 480
tgatgcttcg gatgcacaat ctgcatggga caaagccccc accctcagag ggcagtgatg 540

aagaagaaga ggaggaagat gaagaggatg aagaagagcg gaaacctcag ctggagctgg 600
ccatggtgcc ccactatggt ggcatcaacc gagttcgggt gtcatggctg ggtgaagagc 660
ctgtggctgg ggtgtggtca gagaagggcc aggtggaggt gtttgcgctg cggcggcttc 720
tgcaggtggt ggaggagccc caggccctgg cagccttcct ccgggatgag caggcccaaa 780
tgaagcccat cttctccttc gctggacaca tgggcgaggg ctttgccctt gactggtccc 840
cccgggtgac cggtcgcctg ctgaccggtg actgtcaaaa gaacatccac ctctggacac 900
ctacggacgg cggctcctgg cacgtggacc agcggccatt cgtgggccac acacgctctg 960
tggaggacct gcagtggtca ccgactgaga acacggtgtt tgcctcctgc tcagctgacg 1020
cctccatccg catctgggac atccgggcag cccccagcaa ggcctgcatg ctcaccacag 1080
ccaccgccca tgatggggac gtcaatgtca tcagctggag ccgccgggag cccttcctgc 1140
tcagtggcgg ggatgatggg gccctcaaga tctgggacct tcggcagttc aagtctggtt 1200
ccccagtggc caccttcaag cagcacgtgg cccccgtgac ctccgtcgag tggcaccccc 1260
aggacagcgg ggtctttgca gcctcgggtg cagaccacca gatcacacag tgggacctgg 1320
cagtggagcg ggaccctgag gcgggcgacg tggaggccga ccccggactg gccgacctcc 1380
cgcagcagct gctgttcgtg caccagggcg agaccgagct gaaggagctg cactggcacc 1440
cgcagtgccc agggctcctg gtcagcacgg cgctgtcagg cttcaccatc ttccgcacca 1500
tcagcgtctg aggcgtccca ctggctctga tcttgcttcc tgcttggaaa ctgaagtcga 1560
attgggctcc cctggaaggg gttcattcag gtctgttgac tgagactggc cggcctgtgg 1620
gctgccgtga tggattctgt ttgacgtatt gttctctaga aggcctggct ctgatccagt 1680
gacccctctc accaaagaac tcggtttaac cagggctctg taagaccact cccacccaga 1740
gacttgtgtg gcctggtgtg gcctgtgtgt cggattcctt cctgtcagct gtgacccatt 1800
tgacctgtgt ccccagaacc cagtttttg tttgtttgtt tgagacggag tcttggtctg 1860
tcgcccaggc tggagtgcag tagcacgatc ttggctcact gcaacctccg cctcctgggt 1920
taaagtgatt ctctcagctc agtctcccag gtagctggga ttacaggcat gtgccaccac 1980
accccgttaa tttttgtatt tttagtagag acggggtttc accatgttgg ccaggctggt 2040
ctcaaattct tgatctcaag tgatctgtcc gccccggcct cccagagtgc tgggttggga 2100
ttacaggcgt gagccaccgc gtccggctca ggacccagtt ttggctgctg gttcccagca 2160
ggggactcgg gggatataca gtggctgcac caaattggag gtgtgggttc ctccaacaca 2220
atttgcttct gcccgttgtc ttcctgccag ctgggtttgg ccaggatttc tccgtgtggg 2280
ggctacatgc gaccctctcc cctcctccct gactttagag gctggtgctg tgtcgggagg 2340
aaggtcaggg ctcctgagca gcaataaagg accaggaaga ggcctgaggt gtaaaaaaaa 2400
aaaaaaa 2407

<210> 51
<211> 1158
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 483862CB1

<400> 51
ggaagaccgt cccggatggc ctcggggact gccagtgtgt ggaggtgagc tccgggattg 60
ccggcattcc cgcttctgct ggttgcttca tgctgcaggc tgcggccgtc agccctcgct 120
cgcattggtg gcgctgaggt gccggggcag caagtgacat gtcgtcgggc ctccgcgccg 180
ctgacttccc ccgctggaag cgccacatct cggagcaact gaggcgccgg gaccggctgc 240
agagacaggc gttcgaggag atcatcctgc agtataacaa attgctggaa aagtcagatc 300
ttcattcagt gttggcccag aaactacagg ctgaaaagca tgacgtacca aacaggcacg 360
agataagtcc cggacatgat ggcacatgga atgacaatca gctacaagaa atggcccaac 420
tgaggattaa gcaccaagag gaactgactg aattacacaa gaaacgtggg gagttagctc 480
aactggtgat tgacctgaat aaccaaatgc agcggaagga cagggagatg cagatgaatg 540
aagcaaaaat tgcagaatgt ttgcagacta tctctgacct ggagacggag tgcctagacc 600
tgcgcactaa gctttgtgac cttgaaagag ccaaccagac cctgaaggat gaatatgatg 660

ccctgcagat cactttact gccttggagg gaaaactgag gaaaactacg gaagagaacc 720
aggagctggt caccagatgg atggctgaga aagcccagga agccaatcgg cttaatgcag 780
agaatgaaaa agactccagg aggcggcaag cccggctgca gaaagagctt gcagaagcag 840
caaaggaacc tctaccagtc gaacaggatg atgacattga ggtcattgtg gatgaaactt 900
ctgatcacac agaagagacc tctcctgtgc gagccatcag cagagcagcc acgtaagtag 960
gcaggtttgg gccagggaaa agacagcttg aggagcaata tgaaggcaca tctgtggaca 1020
tgacaaagaa tgcagtcaga tgcacccaac cccttactcc ttttctggga cacccagcgt 1080
cgaacacacc acagaggtgt ctagtctttc tcagttcacc tctgcttaat gggagggaag 1140
cagaacacgg gtggcttc 1158

<210> 52
<211> 1026
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1256777CB1

<400> 52
ctgcgggcct ctctccgtcg ccatggaaac gaaagcggcc aagtagagct ccgtcctgac 60
gcgccgcctc ccgtgggctc cggccggcta agccgcggcg gacaactatg ctgaaagcca 120
agatcctctt cgtggggcct tgcgagagtg gaaaaactgt tttggccaac tttctgacag 180
aatcttctga catcactgaa tacagcccaa cccaaggagt gaggatccta gaatttgaga 240
acccgcatgt taccagcaac aacaaaggca cgggctgtga attcgagcta tgggactgtg 300
gtggcgatgc taagtttgag tcctgctggc cggccctgat gaaggatgct catggagtgg 360
tgatcgtctt caatgctgac atcccaagcc accggaagga aatggagatg tggtattcct 420
gctttgtcca acagccgtcc ttacaggaca cacagtgtat gctaattgca caccacaaac 480
caggctctgg agatgataaa ggaagcctgt ctttgtcgcc acccttgaac aagctgaagc 540
tggtgcactc aaacctggaa gatgaccctg aggagatccg gatggaattc ataaagtatt 600
taaaaagcat aatcaactcc atgtctgaga gcagagacag ggaggagatg tcaattatga 660
cctagccagc cttcacctgg gactgccaca tccccagtga aatcagcatg tttctcggtg 720
cagatctgaa atcacatcca gctcctgatg ttttcttctc cctctgactg cagaggaagt 780
gttcctacct gcaggaaggc acctgtcaca cagggcgttc actcagacca tctgtgctct 840
gccctgagtt cagttgagaa aatcctatta tcaaatttgg atttcctggc cccagaactt 900
cccaaagacc tgtaaaatgg agggatttac cacctcacat atgtccagtt aaacagtttg 960
tggacttgta accgtcgcag cccaatgata caacagtagt ttaatcacgt gaaaaaaaaa 1020
aaaaaa 1026

<210> 53
<211> 2456
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2198779CB1

<400> 53
gacgagcgac gagatgacgg aggagcggtt ggccaacgca tgccggcagt cggtgtaaac 60
aaggcctcgc gccgctgcgg gtcctgcgac cgctcctggc tgggatttcg attggtcggt 120
cagagaggtt acctggaaat ccaacaccgc ccaacacccc tcccgctccc cagtccgggg 180
acttcgatag gattggagaa ggagttgaca ggaggagccc ccgcacagga cctaagaatg 240

ctgtgaccag aagatgggat cgcggaacag cagcagtgca ggatccgggt ccggagaccc 300
ctccgagggc ttgccccgaa gaggggctgg cctgcgtcgg agtgaggaag aggaagaaga 360
ggatgaagat gtggatctgg cccaggtact ggcctatctc ctccgcagag gccaagtgag 420
gttggtgcag ggaggaggtg cagcaaattt acaattcatt caggccctct tggactcaga 480
ggaagagaat gacagagctt gggatggtcg tcttggggat cgatacaacc cacctgtgga 540
tgctacccct gacacccggg agctggaatt caatgagatc aagacacaag tggaactggc 600
cacagggcag ctggggctta ggcgggccgc ccagaagcac agctttcctc gaatgttgca 660
ccagagagaa cggggcctct gccatcgggg aagcttctcc cttggagaac agtctcgagt 720
gatatctcac ttcttgccca atgatctggg cttcactgat agctactctc agaaggcttt 780
ctgtggcatc tacagcaaag atggtcaaat attcatgtct gcttgccaag accagacaat 840
ccgactctat gactgccgat atggccgttt ccgtaaattc aagagcatca aggcccgcga 900
tgtaggctgg agcgtcttgg atgtggcctt cacccctgat gggaaccact tcctctactc 960
tagctggtct gattacattc atatctgcaa tatctatggt gagggagata cacacactgc 1020
cctggatctc aggccagatg agcgtcgctt tgctgtcttc tccattgctg tctcctcaga 1080
tggacgagaa gtactaggag gggccaatga tggctgcctg tatgtctttg accgagaaca 1140
gaaccggcgc acccttcaga ttgagtccca tgaggatgat gtgaatgcag tggcctttgc 1200
tgatataagc tcccaaatcc tgttctctgg gggagatgat gccatctgca aagtgtggga 1260
tcgacgcacc atgcgggagg atgaccccaa gcctgtgggt gcactggctg gacaccagga 1320
tggcatcacc ttcattgaca gcaagggtga tgcccggtat ctgatctcca actctaaaga 1380
ccagaccatc aaactctggg atatccgacg cttttccagc cgggaaggca tggaagcttc 1440
acgccaggct gccacacagc aaaactggga ctatcggtgg cagcaagtgc ccaaaaaagg 1500
gtttactctg catccctacc cagcctggcg gaagctgaag ctcccagggg acagctcctt 1560
gatgacctac cggggccacg gagtgctgca caccctcatc cgctgccggt tctcccccat 1620
tcatagcact ggccagcagt tcatctacag tggctgctcc actggcaaag tggttgtgta 1680
cgaccttcta agtggccaca ttgtgaagaa gctgaccaac cacaaggcct gtgtgcgtga 1740
cgtcagttgg caccccttg aagagaagat tgtcagcagt tcgtgggacg ggaacctgcg 1800
tctgtggcag taccgccagg ctgagtactt ccaggatgac atgccagaat ctgaggaatg 1860
tgccagcgcc cctgccccag tgccccaatc ctctacaccc ttttcctcac cccagtagat 1920
ccaacctcca gccccatata gggtgaacct cttgataagc tctctgcctc ctcctccctt 1980
tctcccttgt ggggaatgtt tggaggaatc actggcattt gatggggaat aacataagcc 2040
tgggctctga gcctcagctg agccctggaa gattctcccc atggggcaga gtggtctcct 2100
tacgtgctca cacccagtca gcttgggtcc ctatctctgg ccagagtttg gcaggactgc 2160
cattatctgg ggtgtggcct ctgccagcaa gagaagtgtc ctgggtgttt ttaatcatgt 2220
ttgaatgtta ggggttggat cctagagtag atgcctgagg ccacatctga acagacctgt 2280
cagccaggcc tgccaggtct tcacgttgag gattcaactg gccaatcaca ggacaggtgt 2340
cctggccttt cttcctgagg tctctagggg aggggcatgg gtaagggtgt ttcctcagca 2400
ccctcctggg gtggggatta tgtctgctgt catgtctggg tctttaaggt aggaca 2456

<210> 54
<211> 1771
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2226116CB1

<400> 54
cggctcgagt taatattttc gttgagcgga accttgctat tccataagag gatgtgtcca 60
gtgttgtgga agatttcatg ttttaaatcc tttgtacaga aatcctgctc ccaagtcaca 120
gataggctga cgggtcagag ggcaagacgt gacccagggc cgagagggtg agtgaccagg 180
aaaatcggat tcatcagttc acttgtttgt ttcagaaacg tgcacaaaga cctgctgcat 240
gaggccctcg tcttcagttt ctgtttcatg cccagcatta aaccaagtat ctcattttgc 300
caatttgact tctgtagggg ccatggcacc tgcaaggtgt ttctcagcaa gattgaggac 360

cgtgtttcag ggcgtggggc attgggcttt gtccacatgg gctggcctga agcccagccg 420
gctactgcca cagcgggctt ctcccaggct gctctcggtc ggccgtgcgg acctcgccaa 480
gcatcaggaa ctcccgggga agaagctgct ctctgagaaa aagctgaaaa ggtactttgt 540
ggactatcgg agagtgcttg tctgtggagg aaacggaggc gctggggcaa gctgcttcca 600
cagtgagccc cgcaaggagt ttggaggccc tgatggaggg gacggaggca acggtggaca 660
cgtcattctg agagttgacc agcaagtcaa gtccctgtcg tcggtcctgt cgcggtacca 720
gggtttcagt ggagaagatg gagggagtaa aaactgcttc gggcgcagtg gcgccgtcct 780
ctacatccgg gtccccgtgg gcacgctggt gaaggaggga ggcagagttg tggccgacct 840
gtcttgcgtg ggagatgagt acattgccgc gctgggcggg gcaggaggga aaggcaaccg 900
cttcttcctg gccaacaaca accgtgcccc tgtgacctgt accccctggac agccaggaca 960
gcagcgagtt ctccacctgg agctcaagac ggtggcccac gccggaatgg tgggattccc 1020
caacgccggg aagtcctcac tgctccgggc catttcaaac gccagacccg ccgtggcttc 1080
ctacccgttc accaccctga agccccacgt cgggatcgtc cactacgaag gccacctaca 1140
aatagcagtg gccgacatcc ccggcatcat acgaggcgcc caccagaaca ggggtctggg 1200
gtccgccttc ctcaggcaca tcgagcgctg ccgctttctc ttgttcgtgg tggatctttc 1260
tcagcctgag ccgtggactc aagttgacga tttaaaatat gaactggaga tgtatgaaaa 1320
gggcctgtct gcgaggcccc acgcaatcgt cgcaaacaag attgacctcc ctgaagccca 1380
agccaatctg tcccagctcc gggatcactt gggacaggag gtcatcgtgc tgtcggcgtt 1440
gaccggcgag aacctggagc agctgctgtt gcacctgaag gtgctgtatg acgcctacgc 1500
ggaggccgag ctgggccagg gccgccagcc gctcaggtgg tagccacgcc agagcggggt 1560
cgcctctggg cctctgtctg agcaaacctg ggtgtgaatt cggtggtttt gaatgcataa 1620
agtgccttgt ggacacgggg gagttgtggt gcttctgggt ctctgggccc cgcctgctgg 1680
cctgagatgc cctcatgttg ggaagcattc cgtgcccccc accccgcctg ccctccgtat 1740
ttcctgcacc tgtcagcctg cgctgactga t 1771

<210> 55
<211> 2724
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2504472CB1

<400> 55
gctgaccagt tggcgacatg gtggcacccg tgctggagac ttctcacgtg ttttgctgcc 60
caaaccgggt gcggggagtc ctgaactgga gctctgggcc cagaggactt ctggcctttg 120
gcacgtcctg ctccgtggtg ctctatgacc ccctgaaaag ggttgttgtt accaacttga 180
atggtcacac cgccccgagtc aattgcatac agtggatttg taaacaggat ggctcccctt 240
ctactgaatt agtttctgga ggatctgata atcaagtgat tcactgggaa atagaggata 300
atcagctttt aaaagcagtg catcttcaag gccatgaagg acctgtttat gcggtgcatg 360
ctgtttacca gaggaggaca tcagatcctg cattatgtac actgatcgtt tctgcagctg 420
cagattctgc tgttcgactc tggtctaaaa agggtccaga agtaatgtgc cttcagactt 480
taaactttgg aaatggattt gctttggctc tctgcttatc ttttttgcca aatactgatg 540
taccaatatt agcatgtggc aatgatgatt gcagaattca catatttgct caacaaaatg 600
atcagtttca gaaagtgctt tctctctgtg gacatgagga ttggattaga ggagtggaat 660
gggcagcctt tggtagagat cttttcctag caagctgttc acaagattgc ctgataagaa 720
tatggaagct gtatataaag tcaacatctt tagaaactca ggatgacgat aacataagac 780
tgaaagaaaa tacttttacc atagaaaatg aaagtgttaa aatagcattt gctgttactc 840
tggagacagt gctagccggt catgaaaact gggtaaatgc agttcactgg caacctgtgt 900
tttacaaaga tggtgtccta cagcagccag tgagattatt atctgcttcc atggataaaa 960
ccatgattct ctgggctcca gatgaagagt caggagtttg gctagaacag gttcgagtag 1020
gtgaagtagg tgggaatact ttgggatttt atgattgcca gttcaatgaa gatggctcca 1080
tgatcattgc tcatgctttc cacggagcgt tgcacctttg gaaacagaat acagttaacc 1140

caagagagtg gactccagag attgtcattt caggacactt tgatggtgtc caagacctag 1200
tctgggatcc agaaggagaa tttattatca ctgttggtac tgatcagaca actagacttt 1260
ttgctccatg gaagagaaaa gaccaatcac aggtgacttg gcatgaaatt gcaaggcctc 1320
agatacatgg gtatgacctg aaatgtttgg caatgattaa tcggtttcag tttgtatctg 1380
gagcagatga aaaagttctt cgggtttttt ctgcacctcg gaattttgtg gaaaattttt 1440
gtgccattac aggacaatca ctgaatcatg tgctctgtaa tcaagatagt gatcttccag 1500
aaggagccac tgtccctgca ttgggattat caaataaagc tgtctttcag ggagatatag 1560
cttctcagcc ttctgatgaa gaggagctgt taactagtac tggttttgag tatcagcagg 1620
tggcctttca gccctccata cttactgagc ctcccactga ggatcatctt ctgcagaata 1680
ctttgtggcc tgaagttcaa aaactatatg ggcacggtta tgaaatattt tgtgttactt 1740
gtaacagttc aaagactctg cttgcctcag cttgtaaggc agctaagaaa gagcatgcag 1800
ctatcattct ttggaacact acatcttgga aacaggtgca gaatttagtt ttccacagtt 1860
tgacagtcac gcagatggcc ttctcaccta atgagaagtt cttactagct gtttccagag 1920
atcgaacctg gtcattgtgg aaaaagcagg atacaatctc acctgagttc gagccagttt 1980
ttagtctttt tgccttcacc aacaaaatta cttctgtgca cagtagaatt atttggtctt 2040
gtgattggag tcctgacagc aagtatttct tcactgggag tcgagacaaa aaggtggttg 2100
tctggggtga gtgcgactcc actgatgact gtattgagca caacattggc cctgctcct 2160
cagtcctgga cgtgggtggg gctgtgacag ctgtcagcgt ctgcccagtg ctccaccctt 2220
ctcaacgata cgtggttgca gtaggattgg agtgtggaaa gatttgctta tatacctgga 2280
aaaagactga tcaagttcca gaaataaatg actggaccca ctgtgtagaa acaagtcaaa 2340
gccaaagtca tacactggct atcagaaaat tatgctggaa gaattgcagt ggaaaaactg 2400
aacagaagga agcagaaggt gctgagtggt tacactttgc aagctgtggt gaagatcaca 2460
ctgtgaagat acacagagtc aataaatgtg cactgtaatg gacttaataa ctacatgctt 2520
gcagtcactg gtatcttaaa atattatcat gtaaacaggt catctttacc ttcataactg 2580
aattgagttt ctgggttttt tttttttttg agatggagtc ttgctttgtc acaacctcca 2640
cctcccaggt tcaagcgatt ctctttcttc agcctcctga gtagctggga ctagaggcac 2700
accaccatgc ccggctaatt tttg 2724

<210> 56
<211> 2963
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3029920CB1

<400> 56
ggccgaagag gctggcaggt ggcgccgtgg ggtgggtgct cctggtgaga ggagtccact 60
ccgtgcgtgc gggcggaggc cggccccega gagccgccga catgaagaaa gacgtgcgga 120
tcctgctggt gggagaacct agagttggga agacatcact gattatgtct ctggtcagtg 180
aagaatttcc agaagaggtt cctccccggg cagaagaaat caccattcca gctgatgtca 240
ccccagagag agttccaaca cacattgtag attactcaga agcagaacag agtgatgaac 300
aacttcatca agaaatatct caggctaatg tcatctgtat agtgtatgcc gttaacaaca 360
agcattctat tgataaggta acaagtcgat ggattcctct cataaatgaa agaacagaca 420
aagacagcag gctgccttta atattggttg ggaacaaatc tgatctggtg gaatatagta 480
gtatggagac catccttcct attatgaacc agtatacaga aatagaaacc tgtgtggagt 540
gttcagcgaa aaacctgaag aacatatcag agctctttta ttacgcacag aaagctgttc 600
ttcatcctac agggcccctg tactgcccag aggagaagga gatgaaacca gcttgtataa 660
aagcccttac tcgtatattt aaaatatctg atcaagataa tgatggtact ctcaatgatg 720
ctgaactcaa cttctttcag aggatttgtt tcaacactcc attagctcct caagctctgg 780
aggatgtcaa gaatgtagtc agaaaacata taagtgatgg tgtggctgac agtgggttga 840
ccctgaaagg ttttctcttt ttacacacac tttttatcca gagagggaga cacgaaacta 900
cttggactgt gcttcgacga tttggttatg atgatgacct ggatttgaca cctgaatatt 960

tgttcccccct gctgaaaata cctcctgatt gcactactga attaaatcat catgcatatt 1020
tatttctcca aagcaccttt gacaagcatg atttggatag agactgtgct ttgtcacctg 1080
atgagcttaa agatttattt aaagttttcc cttacatacc ttgggggcca gatgtgaata 1140
acacagtttg taccaatgaa agaggctgga taacctacca gggattcctt tcccagtgga 1200
cgctcacgac ttatttagat gtacagcggt gcctggaata tttgggctat ctaggctatt 1260
caatattgac tgagcaagag tctcaagctt cagctgttac agtgacaaga gataaaaaga 1320
tagacctgca gaaaaaacaa actcaaagaa atgtgttcag atgtaatgta attggagtga 1380
aaaactgtgg gaaaagtgga gttcttcagg ctcttcttgg aagaaactta atgaggcaga 1440
agaaaattcg tgaagatcat aaatcctact atgcgattaa cactgtttat gtatatggac 1500
aagagaaata cttgttgttg catgatatct cagaatcgga atttctaact gaagctgaaa 1560
tcatttgtga tgttgtatgc ctggtatatg atgtcagcaa tcccaaatcc tttgaatact 1620
gtgccaggat ttttaagcaa cactttatgg acagcagaat accttgctta atcgtagctg 1680
caaagtcaga cctgcatgaa gttaaacaag aatacagtat ttcacctact gatttctgca 1740
ggaaacacaa aatgcctcca ccacaagcct tcacttgcaa tactgctgat gcccccagta 1800
aggatatctt tgttaaattg acaacaatgg ccatgtatcc gcacgtgaca caagctgacc 1860
tcaagagctc cacgttttgg cttcgagcaa gttttggtgc tactgttttt gcagttttgg 1920
gctttgctat gtacaaagca ttattgaaac agcgatgata taaaaagaaa tactgtccct 1980
accaaaaaca aatactttta tgtacattct gaatgcttta agttctgcta gaattattga 2040
gatatttata catgcagagt tactttatta atatttgtaa ttcatgcata agagtatttt 2100
aatgatagtt ataactgcag tattggctag catatggaaa gaaaacagct aacagccaaa 2160
ctaaaatggc taaattccag aggccaaaag ggaatatttt gtaaatatat gtacatattc 2220
aggcaagata tggtctccca agctgagttc tagaaatgat gtttctagac atttctaagt 2280
ggtattgtta gtgctcactt ggctcactct tctaggttta agttagccca gagattgtat 2340
ttactcatgg atcactttat ttatttcaca tttactcaga atgatccttt gggttctata 2400
aggacataag gtacaatttg ccattgtctc tccattttta aaaacataca agtcagtgtc 2460
agcttaccaa catgacattt tttcagtcag ttgtggtagg ccagccttga agccatcgca 2520
cagtctagaa acttgtgtag ctgagtgtgc agctcacctt taagggtgaa gttaggtaaa 2580
agcaattagc agaggcgtta tctatgtgat tatgttgctt ccttgtcagt atgttgaatt 2640
ttatagccct ttcaatgaaa taaaaaaaaa atttgtatat taccaatgtt tttagtttaa 2700
ataaagagtc acccttacta ctgttgaatt tcatcccaag tgtaaatcat tctataatgg 2760
ctgtgtctgt tatagtatat tacagtaact gcatgtgtca ccaagtgttc tatatcaggc 2820
taggataacc tagaggcagt aatttttaa atgataaaat aaatctaatg aatataaact 2880
ctcatgataa acctattttt tccatcatca gccttttcaa gtatttaaat aaataactgc 2940
tgtgtactgt gaaaaaaaaa aaa 2963

<210> 57
<211> 3332
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3332415CB1

<400> 57
gcctggcaga ggctggcggg catcgtgccc gtccctgccg gtctcccggg cacccggcca 60
ccgccccacc ccctcctccc tgccatggag cccgagctgg acgctcagaa gcagcctcga 120
ccgcggaggc gaagccgccg ggcctctggg ctcagcacgg agggagcgac ggggccttcg 180
gccgacacct ccgggtcgga gctggacggg agatgttccc ttcggagagg cagctccttc 240
acattcttaa cacctggccc caactgggac ttcactttga aaagaaaacg cagagagaaa 300
gatgatgatg ttgtaagcct tagcagcctt gatctgaagg agccaagcaa taaaagagtt 360
cgacctctgg ctcgtgtcac gtccttggca aatttaatct ctcctgtaag aaatggagct 420
gtcagacgtt ttggtcaaac aatacagtca tttacccttc gtggtgacca cagatcccca 480
gcctctgccc agaagttttc tagcaggtca acagtcccaa cacccgccaa gagaaggagc 540

agtgcactgt ggtcagagat gctggacatc accatgaagg agtctctcac caccagggag 600
atcagacggc aggaggcaat atatgaaatg tcccgaggtg aacaggattt aattgaggat 660
ctcaaacttg caagaaaggc ctaccatgac cccatgttaa agttgtccat catgtcagaa 720
gaggaactca cacatatatt tggtgatctg gactcttaca tacctctgca tgaagatttg 780
ttgacaagaa taggagaagc aaccaagcct gatggaacag tggagcagat tggtcacatt 840
ctcgtgagct ggttaccgcg cttgaatgcc tacagaggtt actgtagtaa ccagctggca 900
gccaaagctc ttcttgatca aaagaaacag gatccaagag tccaagactt cctccagcga 960
tgtctcgagt ctcccttcag tcgaaaacta gatctttgga gtttcctaga tatccctcga 1020
agtcgcctag tcaaataccc tttactgtta aaagaaattc ttaaacacac tccaaaagag 1080
caccctgatg ttcagcttct ggaggatgct atattgataa tacagggagt cctctctgat 1140
atcaacttga agaaaggtga atccgagtgc cagtattaca tcgacaagct ggagtacctg 1200
gatgaaaagc agagggaccc cagaatcgaa gcgagcaaag tgctgctgtg ccatggggag 1260
ctgcggagca agagtggaca taaactttac atttttcctgt ttcaagacat cttggttctg 1320
actcggcccg tcacacggaa cgaacggcac tcttaccagg tttaccggca gccaatccca 1380
gtccaagagc tagtcctaga agacctgcag gatggagatg tgagaatggg aggctccttt 1440
cgaggagctt tcagtaactc agagaaagct aaaaatatct ttagaattcg cttccatgac 1500
ccctctccag cccagtctca cactctgcaa gccaatgacg tgttccacaa gcagcagtgg 1560
ttcaactgta ttcgagcggc cattgccccc ttccagtcgg caggcagtcc acctgagctg 1620
cagggcctgc cggagctgca cgaagagtgt gaggggaacc acccctctgc gaggaaactc 1680
acagcccaga ggagggcatc cacagtttcc agtgttactc aggtagaagt tgatgaaaac 1740
gcttacagat gtggctctgg catgcagatg gcagaggaca gcaagagctt aaagacacac 1800
cagacacagc ccggcatccg aagagcgagg gacaaagccc tttctggtgg caaacggaaa 1860
gagactttgg tgtagagaag gctctgtgtg ttaactgatg ggagagactg tttgtttata 1920
aatgtgtaca gttttgtttt ctcgtaaggg gagcatcata gggttacttt ataccagttg 1980
taacattttc attgtttttg gttgttcttt tttctttttt taatggcagc taaagatata 2040
cagattactg ttaaattgca gtcctttttt ttttaaagat attttcttga gttatttaga 2100
acatggtaag cctggtattt tttaatcaaa caaaatattt atgaaatggg ttttctctta 2160
attctggatt catcatggct ttctaatacc aattgtaata tttacaatat tcaccaaaac 2220
ttagaatttt gcaaatgctg gaattctgcc agtgtttctt tgctaagcct tgcatgcaaa 2280
atttgaaatt ttaacattgg cacccaaaac ctacatggaa tgtatgtctg gagtatttca 2340
aactttacat tgaaacataa tttccttgga aaacaaacca taagcctgag gaggtttta 2400
tcaactggaa tgctttatat tagtttgttt ttcactgtac attcctcatt ttacattcat 2460
ttaacctgcc gattatttaa ttttttatt gtaaagtagt ttttagcatt tgcttttatt 2520
tttttacttt gatgcctttt caaattggca tgtctttaaa gtatttttct tcctgattaa 2580
aaatgtgtgt gtatgtgtgt gtgtgtgtgt atatatatat attttttaa atcacattaa 2640
ttttaccaag tgaaaccaag ccatactgtt tttgagccaa ttaagaaaat tgccattttt 2700
aaagtgtagc atttcagggt aaagacccat gaaatggctt gatgtattct agactactga 2760
aagaaaacca cttcaaagat tttgttgaaa gttttagtgt tgtctgaaat gcaagaggga 2820
aggtgattgg tagtgagtta aaagaaaaag agaggaaaag agagtagttt tgtcttcaag 2880
taaaatgtct ggttgtgcca gacatttcac aagtgtgaaa ggagatagga gaagctcaac 2940
ttgagggcgt gtagtaagtt gtagaaggct cgagggacg tggacttatt tgccttggtt 3000
tgcaatacct gcaaataatg agtttgaaaa gaaacaatga aatgtgttaa aaatttgacc 3060
atattagata aattttggtg gatttagtca taagatggaa aaagactggt gaatctttta 3120
ttacaaaatg tttctgttaa aatgggatca tcatctttga aaggggggag gaggagtaaa 3180
agcccgatta taatggtgat caattcaagt cagtgttgac tattctgtga aatatatttg 3240
gccagtggaa atgataatca gaaaagactg taaatagatc catccaaatg atttctctgt 3300
acaaatgaat gatactatta aaaaaaaaaa aa 3332

<210> 58
<211> 2617
<212> DNA
<213> Homo sapiens

<220>

<223> Incyte ID No: 4031536CB1

<400> 58
```
tttagtaatg tgcctgtatt acatgtagag agtattcgtc aaccaagagg agttttaaaa  60
tgtcaaaacc gggaaaacct actctaaacc atggcttggt tcctgttgat cttaaaagtg 120
caaaagagcc tctaccacat caaactgtga tgaggatatt tagcattagc atcattgccc 180
aaggcctccc tttttgtcga agacggatga aaagaaagtt ggaccatggt tctgaggtcc 240
gctctttttc tttgggaaag aaaccatgca aagtctcaga atatacaagt accactgggc 300
ttgtaccatg ttcagcaaca ccaacaactt ttggggacct cagagcagcc aatggccaag 360
ggcaacaacg acgccgaatt acatctgtcc agccacctac aggcctccag gaatggctaa 420
aaatgtttca gagctggagt ggaccagaga aattgcttgc tttagatgaa ctcattgata 480
gttgtgaacc aacacaagta aaacatatga tgcaagtgat agaaccccag tttcaacgag 540
acttcatttc attgctccct aaagagttgg cactctatgt gctttcattc ctggaaccca 600
aagacctgct acaagcagct cagacatgtc gctactggag aattttggct gaagacaacc 660
ttctctggag agagaaatgc aaagaagagg ggattgatga accattgcac atcaagagaa 720
gaaaagtaat aaaaccaggt ttcatacaca gtccatggaa aagtgcatac atcagacagc 780
acagaattga tactaactgg aggcgaggag aactcaaatc tcctaaggtg ctgaaaggac 840
atgatgatca tgtgatcaca tgcttacagt tttgtggtaa ccgaatagtt agtggttctg 900
atgacaacac tttaaaagtt tggtcagcag tcacaggcaa atgtctgaga acattagtgg 960
gacatacagg tggagtatgg tcatcacaaa tgagagacaa catcatcatt agtggatcta 1020
cagatcggac actcaaagtg tggaatgcag agactggaga atgtatacac accttatatg 1080
ggcatacttc cactgtgcgt tgtatgcatc ttcatgaaaa aagagttgtt agcggttctc 1140
gagatgccac tcttagggtt tgggatattg agacaggcca gtgtttacat gttttgatgg 1200
gtcatgttgc agcagtccgc tgtgttcaat atgatggcag gagggttgtt agtggagcat 1260
atgattttat ggtaaaggtg tgggatccag agactgaaac ctgtctacac acgttgcagg 1320
ggcatactaa tagagtctat tcattacagt ttgatggtat ccatgtggtg agtggatctc 1380
ttgatacatc aatccgtgtt tgggatgtgg agacagggaa ttgcattcac acgttaacag 1440
ggcaccagtc gttaacaagt ggaatggaac tcaaagacaa tattcttgtc tctgggaatg 1500
cagattctac agttaaaatc tgggatatca aaacaggaca gtgtttacaa acattgcaag 1560
gtcccaacaa gcatcagagt gctgtgacct gtttacagtt caacaagaac tttgtaatta 1620
ccagctcaga tgatggaact gtaaaactat gggacttgaa aacgggtgaa tttattcgaa 1680
acctagtcac attggagagt gggggagtg ggggagttgt gtggcggatc agagcctcaa 1740
acacaaagct ggtgtgtgca gttgggagtc ggaatgggac tgaagaaacc aagctgctgg 1800
tgctggactt tgatgtggac atgaagtgaa gagcagaaaa gatgaatttg tccaattgtg 1860
tagacgatat actccctgcc cttccccctg caaaaagaaa aaaagaaaag aaaaagaaaa 1920
aaatcccttg ttctcagtgg tgcaggatgt tggcttgggg caacagattg aaaagaccta 1980
cagactaaga aggaaaagaa gaagagatga caaaccataa ctgacaagag aggcgtctgc 2040
tgtctcatca cataaaaggc ttcacttttg actgagggca gctttgcaaa atgagacttt 2100
ctaaatcaaa ccaggtgcaa ttatttcttt atttcttct ccagtggtca ttgggcagtg 2160
ttaatgctga aacatcatta cagattctgc tagcctgttc ttttaccact gacagctaga 2220
cacctagaaa ggaactgcaa taatatcaaa acaagtactg gttgactttc taattagaga 2280
gcatctgcaa caaaaagtca tttttctgga gtggaaaagc ttaaaaaaat tactgtgaat 2340
tgtttttgta cagttatcat gaaaagcttt tttttttttt tttgccaacc attgccaatg 2400
tcaatcaatc acagtattag cctctgttaa tctatttact gttgcttcca tatacattct 2460
tcaatgcata tgttgctcaa aggtggcaag ttgtcctggg ttctgtgagt cctgagatgg 2520
atttaattct tgatgctggt gctagaagta ggtcttcaaa tatgggattg ttgtcccaac 2580
cctgtactgt actcccagtg gccaaactta tttatgc                          2617
```